

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 21:56:48 ; Search time 930.45 Seconds

(without alignments)
203.319 Million cell updates/sec

Title: US-09-542-718-1

Perfect score: 106
Sequence: 1 aacggcagcgccctctctgct.....gggcattggcattcattgt 106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pac:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p13:*
10: gb_p14:*
11: gb_p15:*
12: gb_p16:*
13: gb_p17:*
14: gb_p18:*
15: gb_p19:*
16: gb_p20:*
17: gb_p21:*
18: gb_p22:*
19: gb_p23:*
20: gb_p24:*
21: gb_p25:*
22: gb_p26:*
23: gb_p27:*
24: gb_p28:*
25: gb_p29:*
26: gb_p30:*
27: gb_p31:*
28: gb_p32:*
29: gb_p33:*
30: gb_p34:*
31: gb_p35:*
32: gb_p36:*
33: gb_p37:*
34: gb_p38:*
35: gb_p39:*
36: gb_p40:*
37: gb_p41:*
38: gb_p42:*
39: gb_p43:*
40: gb_p44:*
41: gb_p45:*
42: gb_p46:*
43: gb_p47:*

44: gb_hcg7:*
45: em_hcg1:*
46: em_hcg2:*
47: em_hcg3:*
48: em_hcg4:*
49: gb_p13:*
50: gb_p14:*
51: gb_p15:*
52: gb_hcg8:*
53: gb_hcg9:*
54: gb_hcg10:*
55: gb_hcg11:*
56: gb_hcg12:*
57: gb_hcg13:*
58: gb_hcg14:*
59: gb_hcg15:*
60: gb_hcg16:*
61: gb_hcg17:*
62: em_hcg4:*
63: em_hcg5:*
64: em_hcg6:*
65: em_hcg7:*
66: em_hcg8:*
67: gb_hcg18:*
68: gb_hcg19:*
69: gb_hcg20:*
70: gb_hcg21:*
71: gb_hcg22:*
72: gb_hcg23:*
73: gb_hcg24:*
74: gb_hcg25:*
75: gb_hcg26:*
76: gb_hcg27:*
77: gb_hcg28:*
78: gb_hcg29:*
79: gb_hcg30:*
80: gb_hcg31:*
81: gb_v11:*
82: gb_v12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	1242	11 AF022956	AF022956 Homo sapi
2	106	100.0	2305	9 HSBAR	Y00106 Human gene
3	106	100.0	3451	10 HUMADRBR	M5169 Human beta-
4	105.6	99.6	1286	39 AF020305	AF020305 Homo sapi
5	104.4	98.5	1242	11 AF022953	AF022953 Homo sapi
6	104.4	98.5	1242	11 AF022954	AF022954 Homo sapi
7	104.4	98.5	1242	11 AF022955	AF022955 Homo sapi
8	104.4	98.5	1290	39 AF169225	AF169225 Homo sapi
9	104.4	98.5	1290	39 AF203386	AF203386 Homo sapi
10	104.4	98.5	3458	10 HUMADRBR	J02960 Human beta-
11	102.8	97.0	1970	9 HSBAR	X04827 Human mRNA
12	102.8	97.0	132858	43 AC011354	AC011354 Homo sapi
13	102.8	97.0	134419	43 AC011334	AC011334 Homo sapi
14	96.4	90.9	1320	11 MACB2AR	L38905 Macaca mula
15	85.6	78.9	1536	3 AF192345	AF192345 Felis cat
16	83.6	78.9	5288	3 PIGB2AR	AF000134 Sus scrofa
17	80.4	75.8	1959	12 RATADBC	J03024 Rat beta-2
18	80.4	75.8	4190	12 RSB2AR	X13607 Rat beta-2
19	80.4	75.8	4197	12 RATWTA	L39254 Rattus norv
20	80.4	75.8	4928	12 MMB2ARG	X15643 Mouse gene
21	78.8	74.3	2032	3 BTB2ADREC	Z86037 B. laurus mr
22	77.2	72.8	1298	3 CFU73206	U73206 Canis famli
23	77.2	72.8	1948	3 CFB2AR	X94608 C. familiar
24	77.2	72.8	2679	5 A65720	A65720 Sequence 1

25 72.4 68.3 2018 12 MABBR
26 41.1 38.9 3780 12 MABBR
27 35.2 33.2 3399 2 AF0025448
28 33.2 33.2 15851 2 MSU4289
29 33.6 31.7 17581 2 MSU4289
30 33.6 31.7 44484 50 AF2256054
31 33.6 31.7 230177 32 HS3255084
32 33.6 31.7 230177 32 HS3255084
33 31.4 29.6 24149 41 HSAC1076
34 31.4 29.6 24149 41 HSAC1076
35 30 28.3 3710 33 DROMMS
36 30 28.3 47200 42 AC013076
37 30 28.3 65329 43 AC013076
38 30 28.3 113985 51 AC013076
39 30 28.3 139723 51 AC013076
40 30 28.3 185050 57 AC009368
41 30 28.3 230303 34 AC009368
42 30 28.3 291880 34 AC009368
43 29.8 28.1 10348 2 AE003525
44 29.8 28.1 41330 2 SCC24
45 29.2 27.5 19791 43 AC021487
19791 1 SPKRDAD
AC021487
Y10438 Streptomyces

Query Match
Best Local Similarity 100.0%; Score 106; DB 11; Length 1242;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 AACGCGAGCGCTCTCTGTCGACCCCAATAGAGCCATGCGCGAGCAGCAGTACG 60
16 AACGCGAGCGCTCTCTGTCGACCCCAATAGAGCCATGCGCGAGCAGCAGTACG 60
61 CACGCGAGCGCGCTCTCTGTCGACCCCAATAGAGCCATGCGCGAGCAGCAGTACG 75
76 CACGCGAGCGCGCTCTCTGTCGACCCCAATAGAGCCATGCGCGAGCAGCAGTACG 75

ALIGNMENTS

RESULT 1
LOCUS AF022956
DEFINITION Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
ACCESSION AF022956
VERSION AF022956.1
KEYWORDS GI:2570532
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Reihers, E., Innis, M., MacIntyre, N. and Liggett, S.B.
TITLE Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
JOURNAL Am J Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE 93192047
REFERENCE 2 (bases 1 to 1242)
AUTHORS Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
TITLE A polymorphism of the human beta 2-adrenergic receptor
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
JOURNAL J. Biol. Chem. 268 (31), 23116-23121 (1993)
MEDLINE 94043092
REFERENCE 3 (bases 1 to 1242)
AUTHORS Green, S.A., Turk, J., Innis, M. and Liggett, S.B.
TITLE receptor terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
JOURNAL Biochemistry 33 (32), 9414-9419 (1994)
MEDLINE 94347707
REFERENCE 4 (bases 1 to 1242)
AUTHORS Liggett, S.B. and Green, S.A.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave M670564, Cincinnati, OH 45267-0564, USA
FEATURES
Source location/Qualifiers
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADRB2"
1..1242
/codon_start=1
/product="beta2-adrenergic receptor"

BASE COUNT 276 a 330 c 325 g 311 t
ORIGIN
variation
/protein_id="AA82351.1"
/db_xref="GI:2570532"
/translation="MGQPNSSAFLAPNSHAPDHDTQORDEWVWVGIVNSLIV
LAIVGNALVTIAKFERLOTVTNFIYSLACADLVAGLAVPFGAAHILMKKMTG
NMCERFVSTIDVLCVASTLDTVNTFISLACADLVAGLAVPFGAAHILMKKMTG
YSRFQDPIOMHVRATQEAIVNDEYKATISPEYKQSLIKKARVILMKKMTG
INGFETLCMLPEFIVNTVIRFVQWNLQVODRGRHGLRSPKIFKPKATKLG
AFQELICLRSSLSKAGNCTSSNGNTGDSGYHVEKERKMLLCEDLFGTEDFVGHOG
491
/gene="ADRB2"
/note="N-linked glycosylation site"

RESULT 2
LOCUS HSBAR
DEFINITION Human gene for beta-2-adrenergic receptor (beta-2 subtype).
ACCESSION Y00106
VERSION Y00106.1
KEYWORDS beta2-adrenergic receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2305)
AUTHORS Schofield, P.R., Rhee, L.M. and Peralta, E.G.
TITLE Nucleic structure of the human beta-2-adrenergic receptor gene
JOURNAL Nucleic Acids Res. 15 (8), 3636 (1987)
MEDLINE 87203400
REFERENCE 2 (bases 1 to 2305)
AUTHORS Schofield, P.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1987) to the EMBL/GenBank/DBJ databases
1..2305
location/Qualifiers
1..2305
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Manatis human"
194..2305
/note="beta-2-adrenergic receptor (MA 1 - 413)"
/protein_id="U00106"
/db_xref="GI:259371"
/db_xref="CA68289.1"
/translation="MGQPNSSAFLAPNSHAPDHDTQORDEWVWVGIVNSLIV
LAIVGNALVTIAKFERLOTVTNFIYSLACADLVAGLAVPFGAAHILMKKMTG
NMCERFVSTIDVLCVASTLDTVNTFISLACADLVAGLAVPFGAAHILMKKMTG
YSRFQDPIOMHVRATQEAIVNDEYKATISPEYKQSLIKKARVILMKKMTG
INGFETLCMLPEFIVNTVIRFVQWNLQVODRGRHGLRSPKIFKPKATKLG
AFQELICLRSSLSKAGNCTSSNGNTGDSGYHVEKERKMLLCEDLFGTEDFVGHOG
809..817
/note="N-linked glycosylation site"

FEATURES
Source location/Qualifiers
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADRB2"
1..1242
/codon_start=1
/product="beta2-adrenergic receptor"

misc-feature

misc-feature

```

mlsc_feature      836..844
                  /note="N-linked glycosylation site"
mlsc_feature      896..967
                  /note="membrane spanning domain I"
mlsc_feature      1007..1078
                  /note="membrane spanning domain II"
mlsc_feature      1114..1180
                  /note="membrane spanning domain III"
mlsc_feature      1247..1315
                  /note="membrane spanning domain IV"
mlsc_feature      1385..1450
                  /note="membrane spanning domain V"
mlsc_feature      1616..1687
                  /note="membrane spanning domain VI"
mlsc_feature      1712..1774
                  /note="membrane spanning domain VII"
BASE COUNT      495 a      616 c      649 g      545 t
ORIGIN

Query Match      100.0%; Score 106; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 6.2e-17;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 aacgagcagcctctctgtcgtgacccaataagaagcgcgcgcgacgcagcgcagc 60
          |||||||
DB      809 AACGGCAGCGCCTTCTTCTGCGCACCACCAATAGAGCCATCGCGGACGACGATCAGC 868
          |||||||

OY      61 cagcaagaggagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 106
          |||||||
DB      869 CAGCAAGGAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 914
          |||||||

RESULT 3
HUMADRB
LOCUS      HUMADRB      3451 bp      mRNA      PRI      13-FEB-1996
DEFINITION Human beta-2-adrenergic receptor mRNA, complete cds.
ACCESSION M15169 J02728 M16106
VERSION M15169.1 GI:178201
KEYWORDS adrenergic receptor.
SOURCE Homo sapiens (clone: pTF.) (tissue library: Evan Sadler) placenta
CDNA to mRNA.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3451)
          Koblika,B.K., Friele,T., Dohlman,H.G., Bolanowski,M.A.,
          Dixon,R.A., Kellier,P., Caron,M.G. and Lefkowitz,R.J.
          Deletion of the intronless nature of the genes for the human and
          hamster beta 2-adrenergic receptor and their putative promoter
          regions
JOURNAL J Biol Chem. 262 (15), 7321-7327 (1987)
MEDLINE 87222338
REFERENCE 2 (bases 1399 to 1985)
          Koblika,B.K., Dixon,R.A., Friele,T., Dohlman,H.G.,
          Bolanowski,M.A., Sigal,I.S., Yang-Feng,T.L., Francke,U., Caron,M.G.
          and Lefkowitz,R.J.
          CDNA for the human beta 2-adrenergic receptor: a protein with
          multiple membrane-spanning domains and encoded by a gene whose
          chromosomal location is shared with that of the receptor for
          platelet-derived growth factor
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (1), 46-50 (1987)
MEDLINE 87092393
FEATURES
          source
          1..3451
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="pTF."
          /tissue_type="placenta"
          /tissue_lib="Evan Sadler"
          /map="5q31-q32"
          /gene="ADRB2"
          mRNA

```

```

gene      1369..3383
          /note="b-2-adr mRNA (alt.); G00-120-541"
          /gene="ADRB2"
mRNA      1376..3383
          /note="ADRB2"
          /gene="ADRB2"
          /note="b-2-adr mRNA (alt.); G00-120-541"
          1379..3383
          /gene="ADRB2"
          /note="b-2-adr mRNA (alt.); G00-120-541"
          1388..3383
          /gene="ADRB2"
          /note="b-2-adr mRNA (alt.); G00-120-541"
          1487..1546
          /note="b-2-adr mRNA (alt.); G00-120-541"
          /gene="ADRB2"
          /note="putative"
          /codon_start=1
          /db_xref="GDB:G00-120-541"
          /protein_id="AAA8014.1"
          /db_xref="GI:560761"
          /translation="MRPQVRSRPAEPRRGSAR"
          1588..2829
          /gene="ADRB2"
          /codon_start=1
          /db_xref="GDB:G00-120-541"
          /product="beta-2 adrenergic receptor"
          /protein_id="AAA8015.1"
          /db_xref="GI:178202"
          /translation="WGQPGNSAFLLAPNRSHAPDHYTOQDEYVYVGMGIVNSLIV
          LAIVGNLVITAIKFERLQTVTVYFITSACADLVGLAVPGAAHILKMTFG
          NWCFFNSIDVLCVATSIETLVADHYFAITSFQKSLTKNKAIVILNMTIV
          SGLTSELPLOHMYRATNOEAINCYANETCDPEFNQAVAIASIVSFYPLIVVVF
          YSRVPEAKROLOKIDKSGRPNQNLQVSDQDGRGKLRSSKREKHALKTIG
          IIMGFTYCLMDFPIVNYVHYQDMLIREVYIILNIGYVNSGFRNPLTYCSPPRI
          AFQELCLRSLKAVNGYSSNGMTGDSGYHVEQERENKLLCEDLPQTEDFVHGQ
          TVPSNDINSQGRNCSTNDSL"
BASE COUNT      790 a      873 c      895 g      893 t
ORIGIN

Query Match      100.0%; Score 106; DB 10; Length 3451;
Best Local Similarity 100.0%; Pred. No. 5.8e-17;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 aacgagcagcctctctgtcgtgacccaataagaagcgcgcgcgacgcagcgcagc 60
          |||||||
DB      1603 AACGGCAGCGCCTTCTTCTGCGCACCACCAATAGAGCCATCGCGGACGACGATCAGC 1662
          |||||||

OY      61 cagcaagaggagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 106
          |||||||
DB      1663 CAGCAAGGAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1708
          |||||||

RESULT 4
AF202305
LOCUS      AF202305      1286 bp      DNA      PRI      14-DEC-1999
DEFINITION Homo sapiens beta-2 adrenergic receptor gene, complete cds.
ACCESSION AF202305
VERSION AF202305.1 GI:6573152
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1286)
          Rupert,J.L. and Hochachka,P.W.
          Beta-2-adrenergic receptor allele frequencies in two native
          American populations
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1286)
          Rupert,J.L. and Hochachka,P.W.
          Direct Submission
JOURNAL Submitted (04-NOV-1999) Zoology, University of British Columbia,
          6270 University Blvd., Vancouver, BC V6T 1Z4, Canada

```

FEATURES
source
Location/Qualifiers
1. .1286
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q31-q33"
/note="isolated from Quechua speaking Native American"
<35. .>1276
/product="beta-2 adrenergic receptor"
35. .1276
/codon_start=1
/product="beta-2 adrenergic receptor"
/protein_id="AF17569.1"
/db_xref="GI:6573153"
/translation="MGQPGNGSAFLAPNNSHAPDDHVTQQRDEWVVGNGIYMSLIV
LAIVGNVITITAIKFERLQTYTNFTITSLACADLYMGIAVYFGAAHILMKMTFG
NFWCEWTSIDVLCVATSIETLCVIAVDREYFATSPKQSLTKNKAHYITILMWIV
SGLTSLPIQMHWYRATHOEAINCYANETCCDEFTNOAYAIASSIVFYPLIMVIV
YSRVQEAROLQIDKSEGRFHVONLSQVEODGRGHLGRSSKFCLEKRAKTLG
IINGFTLCLPFEIYNIYHVIDNLIRKRYITLMMIGVYNSGFNPLIYCRSPDRI
AFQELICLRSSSLKAYGNGYSNGNTGEOSGHVEQEKENKLLCEDLPETDEYVGHOG
TVPSDNDISGGRNCSTNDLL"

BASE COUNT 282 a 347 c 334 g 320 t 3 others
ORIGIN

Query Match 99.6%; Score 105.6; DB 39; Length 1286;
Best Local Similarity 99.1%; Pred. No. 8.6e-17;
Matches 105; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 aacggcagcgcccttctgtctgacacccaatagaagccatgcgcgcagacacagcgtcacg 60
|||||
DB 50 AACGGCAGCGCCTTCTGTCTGGACCAATGAGACCATGGCGCGGACACGAGCTCAGC 109
|||||
QY 61 cagcaaaaggagcaggt 106
|||||
DB 110 CAGCAAAAGGAGCAGAGT 155
|||||

RESULT 5
AF022953 1242 bp DNA PRI 30-OCT-1997
LOCUS Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION
ACCESSION AF022953.1 GI:2570526
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1242)
TITLE Reihaus, E., Innis, M., Macintyre, N. and Liggett, S.B.
MUTATIONS IN THE GENE ENCODING FOR THE BETA-2-ADRENERGIC RECEPTOR
IN NORMAL AND ASTHMATIC SUBJECTS
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE 93192047
AUTHORS 2 (bases 1 to 1242)
TITLE Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
REFERENCE A polymorphism of the human beta 2-adrenergic receptor within the
AUTHORS fourth transmembrane domain alters ligand binding and functional
TITLE Properties of the receptor
JOURNAL J. Biol. Chem. 268 (31), 23116-23121 (1993)
MEDLINE 94043092
AUTHORS 3 (bases 1 to 1242)
TITLE Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
REFERENCE Amino-terminal polymorphisms of the human beta 2-adrenergic
AUTHORS receptor impart distinct agonist-promoted regulatory properties
TITLE Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL 94347707
REMARK Erratum: [published erratum appears in Biochemistry 1994 Nov
AUTHORS 29; 33(47):14368]]
REFERENCE 4 (bases 1 to 1242)
AUTHORS Liggett, S.B. and Green, S.A.

TITLE Direct Submission
JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave M670564, Cincinnati, OH 45267-0564, USA
FEATURES
source Location/Qualifiers
1. .1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. .1242
/gene="ADRB2"
CDS
1. .1242
/gene="ADRB2"
/codon_start=1
/product="beta2-adrenergic receptor"
/protein_id="AAB82148.1"
/db_xref="GI:2570527"
/translation="MGQPGNGSAFLAPNNSHAPDDHVTQQRDEWVVGNGIYMSLIV
LAIVGNVITITAIKFERLQTYTNFTITSLACADLYMGIAVYFGAAHILMKMTFG
NFWCEWTSIDVLCVATSIETLCVIAVDREYFATSPKQSLTKNKAHYITILMWIV
SGLTSLPIQMHWYRATHOEAINCYANETCCDEFTNOAYAIASSIVFYPLIMVIV
YSRVQEAROLQIDKSEGRFHVONLSQVEODGRGHLGRSSKFCLEKRAKTLG
IINGFTLCLPFEIYNIYHVIDNLIRKRYITLMMIGVYNSGFNPLIYCRSPDRI
AFQELICLRSSSLKAYGNGYSNGNTGEOSGHVEQEKENKLLCEDLPETDEYVGHOG
TVPSDNDISGGRNCSTNDLL"

variation
46
/gene="ADRB2"
/note="Arg16 to Gly polymorphism"
/replace="a"
BASE COUNT 275 a 331 c 326 g 310 t
ORIGIN

Query Match 98.5%; Score 104.4; DB 11; Length 1242;
Best Local Similarity 99.1%; Pred. No. 1.7e-16;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 aacggcagcgcccttctgtctgacacccaatagaagccatgcgcgcagacacagcgtcacg 60
|||||
DB 16 AACGGCAGCGCCTTCTGTCTGGACCAATGAGACCATGGCGCGGACACGAGCTCAGC 75
|||||
QY 61 cagcaaaaggagcaggt 106
|||||
DB 76 CAGCAAAAGGAGCAGAGT 121
|||||

RESULT 6
AF022954 1242 bp DNA PRI 30-OCT-1997
LOCUS Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION
ACCESSION AF022954.1 GI:2570528
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1242)
TITLE Reihaus, E., Innis, M., Macintyre, N. and Liggett, S.B.
MUTATIONS IN THE GENE ENCODING FOR THE BETA-2-ADRENERGIC RECEPTOR
IN NORMAL AND ASTHMATIC SUBJECTS
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE 93192047
AUTHORS 2 (bases 1 to 1242)
TITLE Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
REFERENCE A polymorphism of the human beta 2-adrenergic receptor within the
AUTHORS fourth transmembrane domain alters ligand binding and functional
TITLE Properties of the receptor
JOURNAL J. Biol. Chem. 268 (31), 23116-23121 (1993)
MEDLINE 94043092
AUTHORS 3 (bases 1 to 1242)
TITLE Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
REFERENCE Amino-terminal polymorphisms of the human beta 2-adrenergic
AUTHORS receptor impart distinct agonist-promoted regulatory properties
TITLE Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL 94347707

FEATURES	source	1..1242	/organism="Homo sapiens"	/db_xref="taxon:9606"
gene		1..1242	/gene="ADRB2"	/codon_start=1
CDS		1..1242	/gene="ADRB2"	/product="beta2-adrenergic receptor"
			/protein_id="AA82149.1"	/db_xref="GI:2570529"
			/translation="MGPGNGSAFLAPNRSHADHDVYTERDEVMVYGMIVNSLIV LAIFGVNLVTTALAKFERLQTVNFTTSLACADLVNGLAVPFGAAHILMKMTFG NMFCEPMSIDVLCVTASIERFLVADRYAFAISPEKYGSLTKNNARVILNMVIV SGLSPLPIOMHWRAITHOEAINCYNAMETCODEPTNOAVAIASSVFEYVLVIMV YSRPFQAKNRLOKRIKDSSEGFHQNLSQVQDDRTGHGLRSKFLKEKALKATLG ILKQFTFLCMLPFTIVTVIHIDNLIRKEYITLLNIGVNSFNPILYCRSPFNT AFOLCLLRSSLSKAYGNSYNSNGTGEOSGYHVEOEKENKLLEDLPEDFVGHOG TVPSDNDISOGKNSCTNDSL"	
variation		79	/gene="ADRB2"	/note="Gln27 to Glu polymorphism"
			/replaced="C"	
BASE COUNT	276 a	330 c	326 g	310 t
ORIGIN				
Query Match		98.5%	Score 104.4;	DB 11; Length 1242;
Best Local Similarity		99.1%;	Pred. No. 1.7e-16;	
Matches 105; Conservative		0;	Mismatches 1;	Indels 0; Gaps 0;
Oy	1	aacggcagcgaccttctctgacaccacaaatgaagcacttgccgacgacagactccg	60	
Db	16	AACGGCACCCTTCTCTGCGACCAATGAAACCATGCGCCGACACGACGTCACG	75	
Oy	61	cagcaaggagacagagcttgagtggtggcagggcagcatcgcatgt	106	
Db	76	CAGCAAGGACGACGAGCTGTGGTGGTGGCATGCGCATCGTCATCT	121	
RESULT 7				
AF022955		AF022955	1242 bp	DNA
LOCUS				
DEFINITION				PR1 30-OCT-1997
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
				human.
				human.
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
				Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE				
AUTHORS				1 (bases 1 to 1242)
TITLE				Reithaus, E., Innis, M., Macintyre, N. and Liggett, S.B.
				Mutations in the gene encoding for the beta 2-adrenergic receptor
				in normal and asthmatic subjects
				Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL				
MEDLINE				93192047
REFERENCE				
AUTHORS				2 (bases 1 to 1242)
TITLE				Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
				A polymorphism of the human beta 2-adrenergic receptor within the
				fourth transmembrane domain alters ligand binding and functional
				properties of the receptor
JOURNAL				J. Biol. Chem. 268 (31), 23116-23121 (1993)
MEDLINE				94043092

REFERENCE		3 (bases 1 to 1242)	
AUTHORS		Green,S.A., Turki,J., Innis,M. and Liggett,S.B.	
TITLE		Amino-terminal polymorphisms of the human beta-2-adrenergic receptor impart distinct agonist-promoted regulatory properties	
JOURNAL		Biochemistry 33 (32), 9414-9419 (1994)	
MEDLINE		94347707	
REMARK		Erratum:[[published erratum appears in Biochemistry 1994 Nov 29;33(47):114368]]	
REFERENCE		4 (bases 1 to 1242)	
AUTHORS		Liggett,S.B. and Green,S.A.	
TITLE		Direct Submission	
JOURNAL		Submitted (04-sep-1997) Medicine, Univ of Cincinnati, 231 Bethesda Ave M670564, Cincinnati, OH 45267-0564, USA	
FEATURES		Location/Qualifiers	
SOURCE		1..1242	
gene		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
CDS		1..1242	
		/gene="ADRB2"	
		1..1242	
		/gene="ADRB2"	
		/codon_start=1	
		/product="beta-2-adrenergic receptor"	
		/protein_id="AAB82150.1"	
		/db_xref="GI:2570531"	
		/translaton="MGQPGNGSAFLAPNRSHPADHDVTOORDEVMWGKIVNSLTVLAIVEGNVLVTALAKFERELQTIVNYFITSLACADLVGLAVPFGAAHILMKMTFGNFCEFWTSIDVLCVTASIEITGVIAADVPAITSPEKYOSLLTKNKARVYILMMVIVSGLSFPLDIOMHWRAITHOEA INCYAMETCCDPNTNOAYAIASSIVSPYDLTMVPPYSRFGQNAKRLODKIKDSEGFHQNSLSOVQRQDRTRTGHLRRSKFLKEKKALKLTGILIMOTFTWCLLPFTVINIVHTIONLRKEYIILLNMIGVNSGFNPILTYCRSPDFRIAFOLICLRLSRSLALAYANGVSSNNTGEOSGYHVEOEKENKLCEDELPRTEDFVGHGQITVPSDNIDISOGRCNSTDSL"	
variation		100	
		/gene="ADRB2"	
		/note="Val134 to Met polymorphism"	
		/replace="g"	
BASE COUNT	277 a	331 c	324 g 310 t
ORIGIN			
Query Match		98.5%;	Score 104.4; DB 11; Length 1242;
Best Local Similarity		99.1%;	Pred. No. 1.7e-16;
Matches 105; Conservative 0;		Mismatches 1;	Indels 0; Gaps 0;
Oy	1	aacggcagcgcccttctgtgcgcaccaataagaacatgcccggaccgaagtcaag	60
Dd	16	AACGGCAGCGCCCTTCTGTCTGCACCACAATGAMAACCATTGCCCGCACACACTGACG	75
Oy	61	caggcaaaagagcgaagtgttgagttggtggcatcgggacatcgtaattg	106
Dd	76	CAGCAAAAGGAGCAGAGGTGTGGTGATGGCATGGCATGTCATCT	121
RESULT	8		
LOCUS	AF169225	1290 bp	DNA PRI 10-AUG-1999
DEFINITION	Homo sapiens beta-2-adrenergic receptor gene, complete cds.		
ACCESSION	AF169225		
VERSION	AF169225.1	GI:5714687	
KEYWORDS	.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE			
AUTHORS	1 (bases 1 to 1290)		
TITLE	Rupert,J.R. and Hochachka,P.W.		
JOURNAL	Beta-2-adrenergic receptor allele frequencies in two native		
REFERENCE	unpublished		
AUTHORS	2 (bases 1 to 1290)		
	Rupert,J.R. and Hochachka,P.W.		
TITLE	Direct Submission		

JOURNAL Submitted (14-JUL-1999) Zoology, University of British Columbia,
6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
FEATURES
Location/Qualifiers
Source

1. .1290
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q31-q33"
/cell_type="lymphocyte"
/tissue_type="blood"
/note="isolated from a Quechua speaking native American
heterozygous for a known C/T mutation"
17. .1258
/codon_start=-1
/product="beta-2-adrenergic receptor"
/protein_id="A048036.1"
/db_xref="GI:5714688"
/translation="MGQPGNGSAFLAPNGSHAPDHDTQORDEWVVGMIYMSLIV
LATVGNVITATAKPERLOTYNTFTSLACADLVMLGAVVPEGAHILMKMTFEG
NFWCEFTMSIDVLCVTSIETLCVIANDRFATSPKYSILTKRKARVITLIMWIV
SGLXSEFLPIOMHWYRATHOEAINCYNANETCCDEFTNOAVALASSIVSFVPLVIMVY
YSRVFOAKROLQKIDKSEGRFHYONLSQVEODGRGHGLRRSSKFLKHKALKTIG
IIMGTFTLCWLPPEYIVIVHIDNLIKREYVILLNMGIVNGSFNPLIYCRSPDFRI
AFQELICLRSSLSKATGNGYSNGNTGEGSGYHVEQEKENKLLCEDLPGETDFVGHOG
TVPSDNDISGRCSTNDLSL"

CDS

variation
491
/replace="C/T"
BASE COUNT 287 a 349 c 331 g 322 t 1 others
ORIGIN

Query Match 98.5%; Score 104.4; DB 39; Length 1290;
Best Local Similarity 99.1%; Pred. No. 1.7e-16;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 aacggcagcgcctctctgctgaccccaatagaagcgcacgtgcgcgacacgacgtcacg 60
|||||
Db 32 AACGGCAGCGCCTTCTGCTGGCACCACCAATGAGCCATGCGCGGACGACGACGCTCAG 91
|||||
QY 61 cagcaagggagcagagtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 106
|||||
Db 92 CAGCAAGGAGCAGAGT 137
|||||

RESULT 9
AF203386 1290 bp DNA PRI 28-DEC-1999
LOCUS
DEFINITION Homo sapiens beta-2 adrenergic receptor (ADRB2) gene, complete cds.
ACCESSION AF203386
VERSION AF203386.1 GI:6636495
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1290)
AUTHORS Rupert, J.L. and Hochachka, P.W.
TITLE Beta-2 adrenergic receptor allele frequencies in two Native
American populations
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1290)
AUTHORS Rupert, J.L. and Hochachka, P.W.
TITLE Direct Submission
JOURNAL Submitted (09-NOV-1999) Zoology, University of British Columbia,
6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
FEATURES
Location/Qualifiers
Source

1. .1290
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q32-q34"
/sex="female"
/note="isolated from a Quechuan-speaking Native American"

gene

CDS

26. .1267
/gene="ADRB2"
26. .1267
/gene="ADRB2"
/note="catecholamine receptor"
/codon_start=-1
/product="beta-2 adrenergic receptor"
/protein_id="A048036.1"
/db_xref="GI:6636496"
/translation="MGQPGNGSAFLAPNGSHAPDHDTQORDEWVVGMIYMSLIV
LATVGNVITATAKPERLOTYNTFTSLACADLVMLGAVVPEGAHILMKMTFEG
NFWCEFTMSIDVLCVTSIETLCVIANDRFATSPKYSILTKRKARVITLIMWIV
SGLXSEFLPIOMHWYRATHOEAINCYNANETCCDEFTNOAVALASSIVSFVPLVIMVY
YSRVFOAKROLQKIDKSEGRFHYONLSQVEODGRGHGLRRSSKFLKHKALKTIG
IIMGTFTLCWLPPEYIVIVHIDNLIKREYVILLNMGIVNGSFNPLIYCRSPDFRI
AFQELICLRSSLSKATGNGYSNGNTGEGSGYHVEQEKENKLLCEDLPGETDFVGHOG
TVPSDNDISGRCSTNDLSL"

BASE COUNT 288 a 345 c 333 g 324 t
ORIGIN

Query Match 98.5%; Score 104.4; DB 39; Length 1290;
Best Local Similarity 99.1%; Pred. No. 1.7e-16;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 aacggcagcgcctctctgctgaccccaatagaagcgcacgtgcgcgacacgacgtcacg 60
|||||
Db 41 AACGGCAGCGCCTTCTGCTGGCACCACCAATGAGCCATGCGCGGACGACGACGCTCAG 100
|||||
QY 61 cagcaagggagcagagtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 106
|||||
Db 101 CAGCAAGGAGCAGAGT 146
|||||

RESULT 10
HUMADRBRA
LOCUS
DEFINITION Human beta-2 adrenergic receptor gene, complete cds.
ACCESSION J02960
VERSION J02960.1 GI:178203
KEYWORDS adrenergic receptor; beta-2 adrenergic receptor.
SOURCE Homo sapiens (clone: H-beta-R-[9,10,11]) epidermis DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 3458)
AUTHORS Emorine, L.J., Marullo, S., Delavier, K-Lutcho, C., Kaveri, S.V.,
Druce, T., Trautman, O. and Strosberg, A.D.
TITLE Structure of the gene for human beta 2-adrenergic receptor:
expression and promoter characterization
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (20), 6995-6999 (1987)
MEDLINE 88041037
COMMENT Draft entry and computer-readable copy of sequence [1] kindly
provided by L.J. Emorine, 25-AUG-1987.
FEATURES
Location/Qualifiers
Source

1. .3458
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="H-beta-R-[9,10,11]"
/cell_line="A431"
/tissue_type="epidermis"
/map="5q31-q32"
277. .1032
/note="ORF, putative"
/codon_start=-1
/product="unknown protein"
/protein_id="A048036.1"
/db_xref="GI:560762"
/translation="MFEREYTGLPQVCEGSIISARVQVSTQMTSVSVSLMPPS
ORVFTFCVCHVEFLGASVYSGRVSLDRGDPVPGFCVRAVAVHGLGCVSY
SMAYVRYKSEHVGCGVFPVPCACIGGHSRFLPNVGCRCALCIETSSRAGAGQROYA
ATERPKAPGLAGKRTTSFSLGAPRAGQWMPALDAGVPRGCGQENEGSGRGK
GECPLAPRLPACHMPKVPVPHHGGSSPKVLCI"

mrna 1045. .3057
/note="beta-2-adrenergic receptor mRNA (alt.)."
/db_xref="GI:29373"
/translation="MGQPGNGSAFLIAPRGSHPDHDTQDERVWVVGMIIVSLIV
LAIVGNNVIVITAIKFERLQVTNFFITSLACADLWMLAVPGAAHILKMTFG
NMFCEPMTSIDVLCYIASIETLCVIAVDHYFAITSPFKQSILTKKAVIILMWIV
SGTSEFLPIOMHMYRATHOEAINCYANETCCDFPNOAVAIASISVSPVIVWV
YSRVOEAKROLOKIDKSGRPHYONLSQVEDDGRGHLRRSSRFLCKREHAKLTIG
IIMGFTTLCMLPFTIVNIVHYIDNLIKREVIILNMIGTVNSGRNPLIYCSPPRI
AFQELLCRRSSLKAVNGYSSNGMTGDSGYHVEQENKLLCEDLPETEDFVGHOG
TVPSNDIDSGRNCSTNDSL."

gene 1264. .2505
/gene="ADRB2"
/db_xref="ADRB2"
/codon_start=1
/db_xref="CDB:G00-120-541"
/product="beta-2 adrenergic receptor"
/protein_id="AAA88017.1"
/db_xref="GI:178204"
/translation="MGQPGNGSAFLIAPRGSHPDHDTQDERVWVVGMIIVSLIV
LAIVGNNVIVITAIKFERLQVTNFFITSLACADLWMLAVPGAAHILKMTFG
NMFCEPMTSIDVLCYIASIETLCVIAVDHYFAITSPFKQSILTKKAVIILMWIV
SGTSEFLPIOMHMYRATHOEAINCYANETCCDFPNOAVAIASISVSPVIVWV
YSRVOEAKROLOKIDKSGRPHYONLSQVEDDGRGHLRRSSRFLCKREHAKLTIG
IIMGFTTLCMLPFTIVNIVHYIDNLIKREVIILNMIGTVNSGRNPLIYCSPPRI
AFQELLCRRSSLKAVNGYSSNGMTGDSGYHVEQENKLLCEDLPETEDFVGHOG
TVPSNDIDSGRNCSTNDSL."

CDs 1364. .2505
/gene="ADRB2"
/db_xref="ADRB2"
/codon_start=1
/db_xref="CDB:G00-120-541"
/product="beta-2 adrenergic receptor"
/protein_id="AAA88017.1"
/db_xref="GI:178204"
/translation="MGQPGNGSAFLIAPRGSHPDHDTQDERVWVVGMIIVSLIV
LAIVGNNVIVITAIKFERLQVTNFFITSLACADLWMLAVPGAAHILKMTFG
NMFCEPMTSIDVLCYIASIETLCVIAVDHYFAITSPFKQSILTKKAVIILMWIV
SGTSEFLPIOMHMYRATHOEAINCYANETCCDFPNOAVAIASISVSPVIVWV
YSRVOEAKROLOKIDKSGRPHYONLSQVEDDGRGHLRRSSRFLCKREHAKLTIG
IIMGFTTLCMLPFTIVNIVHYIDNLIKREVIILNMIGTVNSGRNPLIYCSPPRI
AFQELLCRRSSLKAVNGYSSNGMTGDSGYHVEQENKLLCEDLPETEDFVGHOG
TVPSNDIDSGRNCSTNDSL."

BASE COUNT 777 a 890 c 886 g 905 t
ORIGIN 1 bp upstream of EcoRI site; chromosome 5q31-q32.

Query Match 98.5%; Score 104.4; DB 10; Length 3458;
Best Local Similarity 99.1%; Pred. No. 1.4e-16;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 aacggagagcgcctctctgctgcgcacccaatagaagcattgacgcgagacagcagtcacg 60
|||||
Db 1279 AACGGAGCGCCCTTCTTGCGACCCCAATGGAAGCCATGCGCGACCGACGACGTCACG 1338
|||||

Qy 61 cagcaaaaggagagaggt 106
|||||
Db 1339 CAGCAAAAGGAGAGAGGT 1384
|||||

RESULT 11
HSBARR
LOCUS HSBARR 1970 bp mRNA PRI 12-SEP-1993
DEFINITION Human mRNA for brain beta-adrenergic receptor.
ACCESSION X04827
VERSION X04827.1 GI:29372
KEYWORDS beta-adrenergic receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1970)
Chung,F.Z., Lentjes,K.U., Gocayne,J., Fitzgerald,M., Robinson,D.,
Kerlavage,A.R., Fraser,C.M. and Venter,J.C.
Cloning and sequence analysis of the human brain beta-adrenergic
receptor. Evolutionary relationship to rodent and avian
beta-receptors and porcine muscarinic receptors
FEBS Lett. 211 (2), 200-206 (1987)
JOURNAL
MEDLINE
REFERENCE 2 (bases 1 to 1970)
Kerlavage,A.R.
TITLE Direct Submision
AUTHORS Submitted (22-SEP-1987) to the EMBL/Genbank/DBJ databases
JOURNAL Substantial corrections are reported in [2]
COMMENT Data kindly reviewed (22-SEP-1987) by Kerlavage A.R.
FEATURES
source
1. .1970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="neonatal human brain stem"
178. .1419
/note="beta-adrenergic receptor (AA 1-413)"
/codon_start=1

/protein_id="CAA28511.1"
/db_xref="GI:29373"
/db_xref="SWISS-PROT:P07550"
/translation="MGQPGNGSAFLIAPRGSHPDHDTQDERVWVVGMIIVSLIV
LAIVGNNVIVITAIKFERLQVTNFFITSLACADLWMLAVPGAAHILKMTFG
NMFCEPMTSIDVLCYIASIETLCVIAVDHYFAITSPFKQSILTKKAVIILMWIV
SGTSEFLPIOMHMYRATHOEAINCYANETCCDFPNOAVAIASISVSPVIVWV
YSRVOEAKROLOKIDKSGRPHYONLSQVEDDGRGHLRRSSRFLCKREHAKLTIG
IIMGFTTLCMLPFTIVNIVHYIDNLIKREVIILNMIGTVNSGRNPLIYCSPPRI
AFQELLCRRSSLKAVNGYSSNGMTGDSGYHVEQENKLLCEDLPETEDFVGHOG
TVPSNDIDSGRNCSTNDSL."

misc_feature 794. .799
/note="pot. glucocorticoid-responsive element"
965. .970
/note="pot. glucocorticoid-responsive element"
1459. .1464
/note="pot. glucocorticoid-responsive element"
1491. .1496
/note="pot. polyA signal"
1502. .1507
/note="pot. polyA signal"
1952. .1957
/note="pot. polyA signal"
1970
/note="polyA site"

polyA_site
BASE COUNT 459 a 508 c 482 g 521 t
ORIGIN

Query Match 97.0%; Score 102.8; DB 9; Length 1970;
Best Local Similarity 98.1%; Pred. No. 3.9e-16;
Matches 104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 aacggagagcgcctctctgctgcgcacccaatagaagcattgacgcgagacagcagtcacg 60
|||||
Db 193 AACGGAGCGCCCTTCTTGCGACCCCAATGGAAGCCATGCGCGACCGACGACGTCACG 252
|||||

Qy 61 cagcaaaaggagagaggt 106
|||||
Db 253 CAGCAAAAGGAGAGAGGT 298
|||||

RESULT 12
AC011354
LOCUS AC011354 132858 bp DNA HTG 28-JAN-2000
DEFINITION Homo sapiens chromosome 5 clone CIT-HSPC_354F19, WORKING DRAFT
SEQUENCE, 3 ordered pieces.
ACCESSION AC011354
VERSION AC011354.1 GI:6013586
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Granlata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 132858)
DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 132858)
REFERENCE DOE Joint Genome Institute.
TITLE Direct Submision
AUTHORS Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1 1161: contig of 1161 bp in length
1162 60531: contig of 59370 bp in length
60532 132858: contig of 72327 bp in length.
Location/Qualifiers
1. 132858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CIT-HSPC.354F19"
BASE COUNT 40121 a 26539 c 25722 g 40476 t
ORIGIN

Query Match 97.0%; Score 102.8; DB 43; Length 132858;
Best Local Similarity 98.1%; Pred. No. 1.9e-16;
Matches 104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aacggcagcgccttcttctgtgaccccaatagaagccatgcgcgcgagaccagcgtcacg 60
Db 127624 AACGGCAGCGCCTTCTTCTGTGGACCCCAATGAGCCATGCGCGGACACGACGATCAGC 127683

OY 61 cagcaagaagcagagtggt 106
Db 127684 CAGGAAGGAGCAGGT 127729

RESULT 13
AC011334 134419 bp DNA HTG 28-JAN-2000
LOCUS Homo sapiens chromosome 5 clone CIT-HSPC-255N17, WORKING DRAFT
DEFINITION
SEQUENCE, 1 ordered pieces.
AC011334
AC011334.1 GI:6013606
HTG: HTGS_PHASE2; HTGS_DRAFT.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 134419)
DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 134419)
DOE Joint Genome Institute.
REFERENCE Direct Submission
TITLE Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT www.jgi.doe.gov.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 134419: contig of 134419 bp in length.
Location/Qualifiers
1. 134419
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CIT-HSPC-235N17"
BASE COUNT 40270 a 28042 c 27433 g 38674 t
ORIGIN

Query Match 97.0%; Score 102.8; DB 43; Length 134419;
Best Local Similarity 98.1%; Pred. No. 1.9e-16;
Matches 104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aacggcagcgccttcttctgtgaccccaatagaagccatgcgcgcgagaccagcgtcacg 60
Db 54780 AACGGCAGCGCCTTCTTCTGTGGACCCCAATGAGCCATGCGCGGACACGACGATCAGC 54839

QY 61 cagcaagaagcagagtggt 106
Db 54840 CAGGAAGGAGCAGGT 54885

RESULT 14
MACB2AR 1320 bp mRNA PRI 02-OCT-1995
LOCUS Macaca mulatta beta-2 adrenergic receptor (B2AR) mRNA, complete
DEFINITION
cds.
ACCESSION L38905
VERSION L38905.1 GI:1004338
KEYWORDS beta-2 adrenergic receptor.
SOURCE Macaca mulatta cDNA to mRNA.
ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopitheinae;
Macaca.

REFERENCE 1 (bases 1 to 1320)
AUTHORS Amend, A.M. and Guan, X.M.
TITLE Cloning, sequencing, and expression of the rhesus monkey beta 2
adrenergic receptor
JOURNAL DNA Cell Biol. 14 (9), 753-757 (1995)
MEDLINE 95398843
FEATURES
source
Location/Qualifiers
1. 1320
/organism="Macaca mulatta"
/db_xref="taxon:9544"
/note="genomic DNA for this receptor sequenced and found
to be identical with the cDNA"
1. 40
/gene="B2AR"
1. 1320
/gene="B2AR"
1. 1320
/gene="B2AR"
1. 1320
/gene="B2AR"
41. 1288
/gene="B2AR"
/codon_start=1
/product="beta-2 adrenergic receptor"
/protein_id="AAC41914.1"
/translation="MGOPGNSAFILAPNGSHAPDHDVTOERDEANVVGMIYMSLIV
LATVGVNVLITTAIRKRETCYVNFYFITSACADIVMGIAVYFPGAHLIMKMTFG
NFWCEWTSIDVLCVTAISIEILCVIANDRFATLSPKYSILTKKRAVILIMWIV
SGLSFLPIQIMHWYRATHQEAIVARETCDDFTNQAIAIASISYFYPVPIYVYFV
YSRVFOEAKROLOKIDKSEGRFNAOMISQVEDQKRGHGRSRTCLKEHKALKTUG
IMGTFTLCMLPEFIVIVIVIODNLIPKEVYILLMWGVVNSGPNLTYCRSPDRI
AFQELICLRSSIRKACGNGVSSNSNGMTGDSGYHLEQEKENLCEDLPGTEDFVGH
QGTVPNDNISQGRSCSTNDSL"
1289. 1320
/gene="B2AR"
3' UTR

BASE COUNT 289 a 365 c 343 g 323 t
ORIGIN

Query Match 90.9%; Score 96.4; DB 11; Length 1320;
Best Local Similarity 94.3%; Pred. No. 1.6e-14;
Matches 100; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 aacggcagcgccttcttctgtgaccccaatagaagccatgcgcgcgagaccagcgtcacg 60
Db 56 AACGGCAGCGCCTTCTTCTGTGGACCCCAATGAGCCATGCGCGGACACGATCAGC 115

QY 61 cagcaagaagcagagtggt 106
Db 116 CAGGAAGGAGCAGGT 161

```
RESULT 15
AF192345 1536 bp DNA MAM 27-OCT-1999
LOCUS AF192345
DEFINITION Fells catus beta-2 adrenergic receptor gene, complete cds.
ACCESSION AF192345
VERSION AF192345.1 GI:6120128
KEYWORDS
SOURCE
ORGANISM
cat.
Fells catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Fells.
REFERENCE
1 (bases 1 to 1536)
Cully,D.F., Tremml,G. and Zachwieja,S.
Fells domesticus beta adrenergic receptor subtype 2
JOURNAL
Unpublished
2 (bases 1 to 1536)
Cully,D.F., Tremml,G. and Zachwieja,S.
REFERENCE
Submitted (05-OCT-1999) MRL, Merck & Co., Box 2000 R80N-C42,
JOURNAL
Submitted (05-OCT-1999) MRL, Merck & Co., Box 2000 R80N-C42,
Rayway, NJ 07065, USA
FEATURES
source
Location/Qualifiers
1..1536
/organism="Fells catus"
/db_xref="taxon:9685"
/note="synonym: Fells domesticus"
<6..>1262
/product="beta-2 adrenergic receptor"
6..1262
/note="G-protein coupled receptor; beta adrenergic
receptor subtype 2"
/codon_start=1
/product="beta-2 adrenergic receptor"
/db_xref="GI:6120129"
/translacion="MGPGNRSVFLAPNGSHAPDODGTOERNDAVYVGMIVMSLIV
LATVFGNVLVITAIARPERIOTVNTVFTSLACADLVMLGAVPPGASHILKMMTPG
NMWCEPMTSIDVLCYTAETICVAVDRYATSPRYQSLLTKNKARVYLTMTVY
SGLISFLPIOMHMTATHOEATINCTAKETCCDFTNQATINASTIVSFLLPLVAVFV
YSRVQVAQROLOKIDKSEGRPHQNLQVEQDGRSGHGRNASKFCLKENKALKTLG
IMGTFTLCWLPEFTVINIVHVIDNLIPKEVYILNMGVYVNSAFNPPLYCRSPDERI
AFQELCLRRSSLKAYNGYSNNSRTDYAGEHSGPLGQEKDSEVLCEDDPPGTENTL
ANROGTVPNDSIDSGONGSTNDSL"
BASE COUNT 342 a 424 c 398 g 372 t
ORIGIN
```

```
Query Match 78.9%; Score 83.6; DB 3; Length 1536;
Best Local Similarity 86.8%; Pred. No. 2, 2e-11;
Matches 92; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 aacggcagcgccctctctgtgcacccaalagaagccatgcgcgagcaagcagtcacg 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 21 AACCGCAGCGCTCTCTTCTTGGCGCCCAACGAGAACCCAGCGCCGACGAGCGGACG 80
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 cagcaaaaggagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 106
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 81 CAGCAACGAAACGACCGGTGGTGGCGATGGCATGCTCATGT 126
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: September 12, 2000, 23:02:38
Job time: 3950 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

Om nucleic - nucleic search, using sw model

Run on: September 12, 2000, 21:58:58 ; Search time 148.16 Seconds

178.998 million cell updates/sec

Title: US-09-542-718-1

Sequence: 1 aacgacgacgtctctgtct.....gggcatgggcatgcgtcatgt 106

Scoring table:

Gapor 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB Seq Length: 1000000

Post-processing: Minimum match 0%

Listing first 45 summaries

Database : N_Geneseq_36:4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	106	100.0	1999	1	T93250	beta-2 adrenalin r
	2	106	100.0	3451	1	V52614	Human beta-2-adren
	3	94.8	89.4	1400	1	T93249	Beta-2 adrenalin r
	4	77.2	72.8	2679	1	V30468	Canine beta-2 adre
	5	35.2	33.2	9960	1	V58939	Mycobacterium smec
	6	39.4	27.7	5145	1	V45185	Nucleotide sequenc
	7	28.4	26.8	9126	1	T45828	Hepatitis GB virus
	8	28.4	26.8	9126	1	T96447	Fragment of HGBV N
	9	28.4	26.8	9126	1	V39074	Hepatitis GB virus
	C 9	27.8	26.2	3919	1	V10355	Human tollold-1-like
	C 10	27.8	26.2	12019	1	X18867	Alcaligenes sp. Pe
	C 11	27.8	26.2	1608	1	V43679	Multi-specific Bss
	12	27.4	25.8	3321	1	T21872	Human gene signal
	13	27.2	25.7	1419	1	T10928	Nucleotide sequenc
	14	27.2	25.7	1419	1	V62709	Maize male fertill
	15	27.2	25.7	1419	1	V73926	Maize MS45 CDNA. N
	16	27.2	25.7	1419	1	X00462	Zea mays male fert
	17	27.2	25.5	317	1	T76274	Human neutrophil e
	C 18	27	25.5	489	1	X54661	Neutrophil elastase
	C 19	27	25.5	489	1	X54665	Neutrophil elastase
	C 20	27	25.5	1635	1	O24332	Mutant thermostab
	C 21	27	25.5	1635	1	O24338	Mutant thermostab
	22	27	25.5	1686	1	T47961	Mutant taq polymer
	23	27	25.5	1686	1	X27132	Fy4 polymerase cod
	24	27	25.5	1794	1	T32337	Thermus flavus DNA
	25	27	25.5	1794	1	T27255	DNA polymerase I e
	26	27	25.5	1899	1	O24321	Mutant thermostab
	27	27	25.5	1899	1	O24337	Mutant thermostab
	28	27	25.5	2043	1	O24330	Mutant thermostab
	29	27	25.5	2043	1	O24336	Mutant thermostab
	30	27	25.5	2277	1	O24013	Mutant thermostab
	31	27	25.5	2277	1	O24335	Mutant thermostab
	32	27	25.5	2277	1	O24012	Mutant thermostab
	33	27	25.5	2370	1	O24012	Mutant thermostab

34	024324	Mutant thermophil
35	2370	1 T. thermophilus DN
36	080751	T. thermus thermophil
37	25.5	DNA sequence encod
38	2504	1 V63402
39	2505	1 024011
40	024323	Mutant thermostabl
41	2505	1 V53993
42	25.5	Nucleotide sequenc
43	2511	1 V53994
44	25.5	Nucleotide sequenc
45	2511	1 V53995
46	013748	T. thermophilus DN
47	25.5	DNA polymerase I h
48	3048	1 T27254
49	25.5	Thermus flavus DNAs
50	3049	1 T32326

ALIGNMENTS

```

RESULT      1
T93250
ID      T93250 standard; cDNA to mRNA; 1999 BP.
AC      T93250;
DC      20-Apr-1998 (first entry)
DE      Beta-2 adrenalin receptor subtype coding sequence.
KW      Beta-2 adrenalin subtype; cyanopindrol; agonist; antagonist;
        asthmatic disease; as.
OS      Homo sapiens.
FH      Key
FT      Location/Qualifiers
CDS      190..1431
        /*tag= a

MO9735963-A1.
PD      02-OCT-1997.
PE      24-MAR-1997; J00982.
PR      27-MAR-1986; JP-072914.
PA      (DAIN ) DAINIPPON PHARM CO LTD.
PI      Fujii K, Fututani Y, Kawashima H, Nomura A, Yano K;
        WP1: 97-489627/45.
DR      P-PSDB: W43450.
DT      Novel beta-2 adrenalin receptor sub-type - useful for screening for
PT      agonists and antagonists and researching asthmatic diseases
PS      disclosure; Page 27-30; 47pp; Japanese.
CC      This sequence encodes the protein of the invention. The protein of the
CC      invention is a beta-2 adrenalin receptor subtype with kd value of
CC      approximately 75 pM against 125I-cyanopindrol. The protein can be used in
CC      screening for agonists and antagonists, which are useful in researching
CC      asthmatic diseases.
SQ      Sequence 1999 BP; 477 A; 513 C; 485 G; 524 T;

Query Match      100.0%; Score 106; DB 1; Length 1999;
Best Local Similarity 100.0%; Pred. No. 7,3e-23;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY      1 aacgcacgcgcctctctgcctgcacccaatagaagcactgcgcgcgcgcacacgcacgtcaccg 60
        |||||||
DB      205 AACGGCAGCGCCTTCTTCTGCTGGCACCACCAATAGAACGCATGCGCGGACCAACGACGCTCACC 264
        |||||||

OY      61 cagcaaaaggagcagaggtgtgtggtgtgtgtggtgcatgtgcacgtcattgt 106
        |||||||
DB      265 CAGCAAAAGGAGCAGAGGTGTGTGGTGTGTGGCATTGCGCATGCGCATGCTCATGT 310
        |||||||

RESULT      2
V52614
ID      V52614 standard; cDNA; 3451 BP.
AC      V52614;
DC      21-DEC-1998 (first entry)
DE      Human beta-2-adrenergic receptor cDNA.
KW      Beta-2-adrenergic receptor; human; asthma; beta-agonist;
        polymorphism; ds.
OS      Homo sapiens.
FH      Key
FT      Location/Qualifiers
CDS      1588..2829
        /*tag= a

```

```

FT variation: 1633 /tag= b "A to G substitution, results in Arg16
FT /note= "A to G substitution, results in Arg16
FT /note= "A to G substitution, results in Arg16
FT to Gly amino acid change"
PD W09839477-A2.
PD 11-SEP-1998.
PF 26-FEB-1998; U03908.
PR 03-MAR-1997; US-811441.
PA (BOHM ) BRIGHAM & WOMENS HOSPITAL.
PI Boushey H, Chinchilli VM, Drazen JM, Fish JE, Ford JG,
PI Martin RJ.
DR WP: 98-506372/43.
DR P-SDB: W75777.
PT Diagnosing asthma patients predisposed to adverse beta-agonist
PT reactions upon regular administration - by identifying patients
PT homozygous for allele encoding Arg at position 16 of
PS beta2-adrenergic receptor protein.
PS This cDNA sequence codes for human beta-2-adrenergic receptor (see
PS W75777) having an arginine residue at position 16. A novel method
CC for identifying individuals susceptible to adverse responses to
CC regular administration of beta-agonists comprises: (a) identifying
CC in a genomic nucleic acid sample from the individual first and
CC second alleles of the beta 2-adrenergic receptor gene, and (b)
CC classifying an individual as susceptible if first and second
CC alleles both encode Arg at residue 16 of the beta 2-adrenergic
CC receptor protein. Beta 2-adrenergic receptor gene alleles may be
CC identified by any known method e.g. denaturing gel electrophoresis
CC or PCR amplification (see also V52615-17). Identification
CC preferably comprises amplifying a portion of each allele which
CC includes the sequence encoding residue 16, and optionally also
CC comprises determining nucleotide sequences of these portions (e.g.
CC by automated sequence analysis). The invention identifies a known
CC polymorphism in the beta 2-adrenergic receptor gene as being linked
CC to adverse responses to regular beta-agonist administration;
CC position 16 of the encoded protein can be either Arg or Gly, and
CC individuals homozygous for Arg16 are more susceptible. 893 T;
SQ Sequence 3451 BP; 790 A; 873 C; 895 G; 893 T;

Query Match 100.0%; Score 106; DB 1; Length 3451;
Best Local Similarity 100.0%; Pred No. 7.8e-23;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aacggcagcgcccttcgtgcgcaccaccaatagaagccatgcccgcgaacgaagctcacg 60
DB 1603 AACGGCAGCGCGCTTCTTGTCGGCGACCAATAGAAGCCATCGCGCGCACGACGTACAG 1662
QY 61 cagcaaaaggagcagagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 106
DB 1663 CAGCAAAGGCAGAGGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1708

RESULT 3
TS3249
ID TS3249 standard; cDNA to mRNA: 1400 BP.
AC TS3249;
DT 20-APR-1998 (first entry)
DE Beta-2 adrenalin receptor subtype coding sequence.
KW Beta-2 adrenalin subtype; cyanopindrol; agonist; antagonist;
KM asthmatic disease; ss.
OS Homo sapiens.
PI Key location/Qualifiers
PH 101. 1348
FT CDS /tag= a
FT /transl_except= (pos: 338..340, aa: Xaa)
FT /note= "Xaa-Leu, Ile"

W09735963-A1.
PD 02-OCT-1997.
PF 24-MAR-1997; J00892.
PR 27-MAR-1996; JP-072914.
PA (DAIN ) DAINIPPON PHARM CO LTD.
PI Fujii K, Furutani Y, Kawashima H, Nomura A, Yano K;
```

WP1: 97-489627/45.
DR p-PDB: W34319.
PT Novel beta-2 adrenalin receptor sub-type - useful for screening for agonists and antagonists and researching asthmatic diseases
PS Claim 5; Page 23-26; 47pp. Japanese.
CC This sequence encodes the protein of the invention. The protein of the invention is a beta-2 adrenalin receptor subtype with Kd value of approximately 75 pM against 125I-cyanopindolol. The protein can be used in screening for agonists and antagonists, which are useful in researching asthmatic diseases.

SC Sequence 1400 BP; 304 A; 402 C; 363 G; 328 T;

Query Match 89.4%; Score 94.8; DB 1; Length 1400;
Best Local Similarity 93.4%; Pred. No. 1.4e-19;
Matches 99; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 aacggcagcgcttctgtgaccaccaatagaagccatgcgcggacacgaagtcacg 60
|||||
DB 116 AACGGCAGCGCTTCTGTGACGCCAACGGAAGGCATCGCGGACACGACGACTAAG 175
|||||

OY 61 cacaagaaggacgagtggtggtggtggtggtggtggtggtggtggtggtggt 106
|||||
DB 176 CAGGAACGGGACGAGGCGTGGTGTGTGGCATGGGCATTCGTCAATG 221
|||||

RESULT 4
V30468 V30468 standard; CDNA to mRNA; 2679 BP.
ID V30468; (first entry)
DE 14-OCT-1998
DT Canine beta-2 adrenergic receptor coding sequence.
KM Canine; beta-adrenergic receptor; brown adipose tissue; probe: human;
KW hybridisation; ligand; ss.
OS Canis familiaris.
FH Key Location/Qualifiers
FT key 169..1416
FT CDS /product= "beta-2 adrenergic receptor"
FN MO9735973-A2.
PN 02-OCT-1997.
PF 26-MAR-1997; F00537.
PI 26-MAR-1996; FR-003730.
PR (VEIT-) VERIGEN.
PA Drumare MF, Lenzon G, Pietri-Rouxel F, Strosberg AD;
PI WP1; 98-032136/03.
DR Canine beta 2 and beta 3 adrenergic receptors and coding sequences -
P-PDBS: W44932.
PT useful for identifying specific ligands and (ant)agonists to develop
specific treatments for obesity in dogs
PS Claim 1; Page 45-46; 79pp. French.
CC This sequence represent the coding region of the canine beta 2-adrenergic
receptor (RA-Ca-b2) gene. The sequence was isolated from a cDNA library
constructed from polyA+ RNA purified from dog brown adipose tissue cells.
CC The probe was a 600 bp fragment of the coding region of the human beta-3
adrenergic receptor covering the region from the initiation codon to
transmembrane domain 5 (TM5). The full length insert was cloned into M13
for sequencing using primers V30491-V30510. The sequence can then be
expressed e.g. in a mammalian cell, by subcloning into an expression
vector such as pCDNA3. The beta-2 receptor can be used in comparative
structure-function studies, e.g. for differential screening of ligands
specific for RA-Ca-b2 or RA-Ca-b3 (W44933).
CC Sequence 2679 BP; 577 A; 736 C; 724 G; 642 T;

Query Match 72.8%; Score 77.2; DB 1; Length 2679;
Best Local Similarity 83.0%; Pred. No. 2.2e-14;
Matches 88; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 1 aacggaagccttctgtgacccaataagaagccatgcccggaccacgaagtcacg 60
|||||
DB 184 AACCGAGCGCTTCTGTGCGCCCAACGAGGACCACGCGCCGACGAGGAGACTCG 243
|||||

Page 3

FT	CDS	252..3293
FT		/*tag=
FT		/product="Human cardiac/brain tolloid-like protein"

```

FE 23-DEC-1997: 310521.
PR 02-JAN-1997: US-034471.
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PI Arieth AD, Elshourbagy MA, Li X, Willette RN;
DR WPI: 98-379062/33.
P-PSDB: W61539.
PR New nucleic acid encoding human cardiac/brain tollold-like protein -
PT and related products, used for treating and diagnosing, e.g.
PS restenosis, atherosclerosis and congestive heart failure
PS Claim 2: Page 5-10; 34pp; English.
CC The human cardiac/brain tollold-like protein (hc/BrLP) inadequate or
CC excessive activity is associated with restenosis, atherosclerosis,
CC congestive heart failure, chronic obstructive pulmonary disease, benign
CC prostatic hypertrophy, glomerulonephritis, nephritis, fibrosis, gliosis,
CC cirrhosis and anomalous wound healing (e.g. keloids). Disease or
CC susceptibility to disease related to expression or activity of hc/BrLP
CC can be diagnosed by detecting a mutation in the gene encoding of hc/BrLP
CC hc/BrLP. Polypeptides and polynucleotides of hc/BrLP can be used in the
CC treatment of the above diseases.
SQ Sequence 5145 BP: 1536 A; 992 C; 1179 G; 1438 T;

Query Match 27.7%: Score 29.4; DB 1; Length 5145;
Best Local Similarity 56.88; Pred. No. 2.7;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 6 cagcgctctctgctcgtgacccaatagaagccatcgccgcgagccagcagctcagcgaca 65
   |||||
Db 260 CAACCCCATCTCTCCCTCCGCCGAGAAAGGAAGGAGCGGACCCAGAAATTAGACAGCC 201
   |||||

QY 6 aaggagcagaggtgctggtgctggtgagcagtcgcatcg 100
   |||||
Db 200 GCGGTCCGCGCTGATGTGCGGGGCTTAGGGGACCG 166
   |||||

RESULT 7
T45828/C T45828 standard; DNA: 9126 BP.
AC T45828;
DT 10-MAR-1997 (first entry)
DE Hepatitis GB virus isolate C large open reading frame.
KW Hepatitis GB virus; HGBV; diagnosis; vaccine; ss.
OS Hepatitis GB virus isolate C genotype 1.
FH Key Location/Qualifiers
   CDS 344..9065
      /*tag- a
      /*transl_except= 2280..2282;_aa:Val
FT misc_difference 356..357
      /*tag- b
      /*note- "bases 356-357 are superfluous to the
FT coding sequence and should be deleted to
FT maintain the reading frame for the
FT protein sequence provided in the
FT specification"
      494
FT misc_difference 494
      /*tag- c
      /*note- "base 494 is superfluous to the coding
FT sequence and should be deleted to
FT maintain the reading frame for the
FT protein sequence provided in the
FT specification"
      569
FT misc_difference 569
      /*tag- d
      /*note- "base 569 is superfluous to the coding
FT sequence and should be deleted to
FT maintain the reading frame for the
FT protein sequence provided in the
FT specification"

```


THIS PAGE BLANK (USPTO)

Accession	Sequence	Length
QY	1 aacggcagagccctctcttggtgaccacaaatgaagccatgtgcccggaccacagatccag	60
Db	16 AACGGCAGGCGCTCTTGTGTCGACACCCAAATGAAGCCATGGCGCGGACACAGAGTCAG	75
QY	61 caaccaaaggacagagtctgtggtgtgtggtcattggtgcattgcctgcatgtc 106	
Db	76 CACCAAAAGGACAGGTGTGGTGTGTGTGGGCAATGGGCATCGTCATGT 121	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match          57.7%; Score 61.2; DB 6; Length 1254;
Best Local Similarity 80.0%; Pred. No. 1.9e-10;
Matches 72; Conservative 0; Mismatches 18; Indels 0; Gaps
QY 1 aacgacagcgcctctcttgctgaccccaatagaagcatalgcgcgacacacgacgtcaag 60
    |||||
Db 16 AACGACAGTACTTTCTTGCTGCACCAACCAAGCAACGCATGTGCCAGACCAAGATGTCACT 75
    |||||
QY 61 cagcaaaaggacagagtgctggtgtgtgggc 90
    |||||
Db 76 GAGGAAAGGGGACGACAAACGATGGGTGTAAGGC 105
    |||||

RESULT 3
US-08-822-586-46
; Sequence 46; Application US/08822586
; Patent No. 6015890

```

GENERAL INFORMATION:
APPLICANT: WILLIAM R. JACOBS, JR., JAMES M. MUSSER AND
APPLICANT: AMALIO TELENITI
TITLE OF INVENTION: AN EMB3AB OPERON OF MYCOBACTERIA AND
TITLE OF INVENTION: MUTANTS THEREOF
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
STREET: 90 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE
MEDIUM TYPE: DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,586
FILING DATE: MARCH 20, 1997
ATTORNEY/AGENT INFORMATION:
NAME: ELIZABETH A. BOGOSIAN
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 96700/437
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
TELEX: TWX 710-581-4766
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 9960
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
US-08-822-586-46

	Query Match	33.2%;	Score 35.2;	DB 5;	Length 9960;	
	Best Local Similarity	58.7%;	Pred. No. 0.022;			
	Matches	61;	Conservative	0;	Mismatches 43;	Indels 0; Gaps 0;
QY	2	aaggcagcgcttcttgctgtagccccaatagaagcatcagcgacggcagcacgcgtcaagc	61			
Db	9665	ACGCCAACGGCGTCAACCAGAGTGCCCAAGTTCCGATCTCGCCSACTRACGCCAAGC	9724			
QY	62	agcaaaaggacgaagtgtgagtgtgagcatgpggacatgcgtcatg	105			
Db	9725	TGCAGAGCACCGACACGTGGCAGGCGCATTCACAGCGGGCCTTG	9768			

RESULT 4
US-08-991-408-1/c
Sequence 1, Application US/08991408
Patent No. 6008017
GENERAL INFORMATION:
APPLICANT: ARLETH, ANTHONY J.
APPLICANT: WILLETTE, ROBERT N.
APPLICANT: ELSHOURBAGY, NABIL A.
APPLICANT: LI, XIANGONG
TITLE OF INVENTION: HUMAN CARDIAC/BRAIN TOLLOID-LIKE
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

```

Query Match Similarity      27.7%: Score 29.4; DB 5; Length 5145:
Best Local Similarity      56.8%; Pred No. 1.1;
Matches    54: Conservative   0; Mismatches   41; Indels     0; Gaps     0

Oy      6 cagcgcttctgtcgtgcaaccataagaagccatgcgccgaacacgactcagcaga 65
        || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      260 CAACCCCATCTCTCCTCCCCCGAGGAAAGAAGGACGGACCCAGAAATTATGCAACC 201

Oy      66 aaaggacgaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 100
        || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      200 GCCGTCCGCCCTGATGTCCGGGGCTAAGGGGCACCG 166

RESULT      5
US-08-580-038-26/C
Sequence 26, Application US/08580038
Patent No. 5807670

GENERAL INFORMATION:
APPLICANT: Muertoff, A. S.
APPLICANT: Simons, J. N.
APPLICANT: Birkenmeyer, L.
APPLICANT: Leary, T. P.
APPLICANT: Erker, J. C.
APPLICANT: Dosai, S. M.
APPLICANT: Mushahwar, I. K.
APPLICANT: Chalmers, M.
APPLICANT: Dawson, G. J.
TITLE OF INVENTION: Detection of Hepatitis GB Virus
TITLE OF INVENTION: Genotypes
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

Query Match	25.8%	Score 28.4	DB 2	Length 9126
Best Local Similarity	60.3%	Pred. NO.2.5		
Matches 47	Conservative	0	Mismatches 31	Indels 0
			Gaps	0
0Y	3	cgagagcgccctctctgtcgcacccaatagagccatgcccgcagccagcgtcagca	62	
Db	1152	CGGGCGCCCTTGTACATGCCCCGCAATGATGCACGAGGAGAACCAATGACATAATCGCT	1093	
0Y	63	gcaaaaggagcagagtgtg	80	
Db	1092	GCTCCAGCAGGAGGAGGG	1075	

```

1      RESULT      6
2      US-08-639-857-3/c
3      Sequence 3, Application US/08639857
4      Patent No. 5955318
5      GENERAL INFORMATION:
6      APPLICANT: Simons, J. N.
7      APPLICANT: Desai, S. M.
8      APPLICANT: Mushawar, I. K.
9      TITLE OF INVENTION: TRANSGENS AND METHODS USEFUL FOR CONTROLLING THE
10     TITLE OF INVENTION: TRANSLATION OF HEPATITIS GB PROTEINS
11     NUMBER OF SEQUENCES: 32
12     CORRESPONDENCE ADDRESSES:
13     ADDRESSEE: Abbott Laboratories
14     STREET: 100 Abbott Park Rd
15     CITY: Abbott Park
16     STATE: IL
17     COUNTRY: USA
18     ZIP: 60064
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: Patentin Release #1.0, Version #1.30
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/639,857
26     FILING DATE:
27     CLASSIFICATION: 435
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Porembski, Priscilla E.
30     REGISTRATION NUMBER: 33,207
31     REFERENCE/DOCKET NUMBER: 5793_US.P1
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: 708-937-0378
34     TELEFAX: 708-938-2623
35     INFORMATION FOR SEQ ID NO: 3:
36     SEQUENCE CHARACTERISTICS:
37     LENGTH: 9126 base pairs
38     TYPE: nucleic acid
39     STRANDEDNESS: single
40     TOPOLOGY: linear
41

```

[illegible]

US-08-815-688A-2

US-08-815-688A-2

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: U.S.
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,739B
FILING DATE: 02-AUG-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/537,183
FILING DATE: 12-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0125R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248 4897
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1419 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-103-739B-1

Query Match 25.7%; Score 27.2; DB 1; Length 1419;
Best Local Similarity 67.9%; Pred. No. 4.3;
Matches 38; Conservative 0; Mismatches 18; Indels 0; Caps 0;

Db 1143 CGAGGACCGGGGCCACGAGTGATGAACTGCTGACGCGAGTGCGGAGTGCGCA 1198

RESULT 11
US-08-474-404-1
Sequence 1, Application US/08474404
Patent No. 5824524
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, Marc C.
APPLICANT: BEACH, Larry R.
APPLICANT: HOWARD, John A.
APPLICANT: HUFMAN, Gary A.
TITLE OF INVENTION: Nucleotide Sequences Mediating Fertility and Method of Using
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: U.S.
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,404
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:

DY 36 ccaatgcgcccgcacacagcctccacgcagaaggaacggagtgttgcgtcggtgcga 91
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1143 CGAGGACCGGGGCACAGAGGTCTATCAAGCTCTGTTCAGCCGAGTCCGGGAGAGTGCGCA 1198

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic : nucleic search, using sw model

Run on: September 12, 2000, 21:43:38 : Search time 1893.64 Seconds
(without alignments)
246.848 Million cell updates/sec

Title: US-09-542-718-1

Perfect score: 106
Sequence: 1 aaagcagcagcctctctgtc.....gggcagtgagcagtcgcatgt 106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_est56:*
94: gb_est57:*
95: gb_est58:*
96: gb_est59:*
97: em_gsa1:*
98: em_gsa2:*
99: em_gsa3:*
100: em_gsa4:*
101: gb_gsa5:*
102: gb_gsa6:*
103: gb_gsa7:*
104: gb_gsa8:*
105: gb_gsa9:*
106: em_gsa5:*
107: em_gsa6:*
108: em_gsa7:*
109: em_gsa8:*
110: em_gsa9:*
111: em_gsa10:*
112: em_gsa11:*
113: gb_gsa10:*
114: gb_gsa11:*
115: em_gsa12:*
116: gb_gsa12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	81.2	76.6	377	80	C18250	C18250 C18250 Huma
2	80.4	75.8	744	33	AA980445	AA980445 ua41d02.r
3	32.6	31.7	165	60	AV306064	AV306064 AV306064
4	32.8	30.9	255	22	AA209561	AA209561 mu71h05.r
5	32.8	30.9	286	42	AI661315	AI661315 va15f07.x
6	32.8	30.9	370	27	AA470234	AA470234 vq90h08.r
7	32	30.2	172	49	AV055159	AV055159 AV055159
8	31.4	29.6	233	23	AA254331	AA254331 va15f08.r
9	31.2	29.4	178	50	AV129211	AV129211 AV129211
10	31.2	29.4	200	60	AV298968	AV298968 AV298968
11	31.2	29.4	243	60	AV297381	AV297381 AV297381
12	29.6	27.9	187	49	AV107772	AV107772 AV107772
13	29.6	27.9	223	61	AV340114	AV340114 AV340114
14	29.6	27.9	305	45	AI838527	AI838527 ut-m-10-
15	29.6	27.9	354	28	AA607055	AA607055 vns3e04.r
16	29.6	27.9	854	116	AO875048	AO875048 v120f11.m
17	29.2	27.5	150	61	AV352055	AV352055 AV352055
18	29.2	27.5	310	50	AV138058	AV138058 AV138058
19	29.2	27.5	567	69	AM225456	AM225456 SWD25CAU
20	29.2	27.5	791	116	AO858041	AO858041 dbe0011n
21	29	27.4	125	27	AA515527	AA515527 nfc6912.s
22	29	27.4	156	60	AV302729	AV302729 AV302729
23	29	27.4	362	33	AA980400	AA980400 ua52g10.r
24	29	27.4	477	47	AI135598	AI135598 DKFZP762E
25	29	27.4	172	60	AV323541	AV323541 AV323541
26	28.8	27.2	254	33	AA980836	AA980836 ua46c02.r
27	28.8	27.2	318	34	AI021469	AI021469 ub08f07.r
28	28.8	27.2	438	28	AA616441	AA616441 v008f04.r
29	28.8	27.2	620	69	AA185521	AA185521 se80h01.y
30	28.8	27.2	646	26	AA458282	AA458282 vq48a11.r
31	28.6	27.0	403	89	TS4325	TS4325 ya91d10.s3
32	28.6	26.8	155	45	AI843984	AI843984 ut-m-AD1-
33	28.4	26.8	450	48	AU063749	AU063749 AU063749
34	28.4	26.8	525	48	AU063240	AU063240 AU063240
35	28.4	26.8	455	38	AI347556	AI347556 g098g12.x
36	28.2	26.6	178	50	AV157571	AV157571 AV157571
37	28	26.4	321	80	AA708048	AA708048 ab81l1n.r
38	28	26.4	406	38	AI329823	AI329823 bf1l1n.r
39	28	26.4	412	79	AA671056	AA671056 LG1_284_E
40	27.8	26.2	407	23	AA278503	AA278503 zs81e04.r
41	27.8	26.2	408	43	AI706391	AI706391 ut-r-AE1-
42	27.8	26.2	419	80	C29133	C29133 C29133
43	27.8	26.2	421	22	AA218070	AA218070 mw01g03.r
44	27.8	26.2	425	81	C93520	C93520 C93520
45	27.8	26.2	482	70	AM252709	AM252709 ut-r-BD0-

ALIGNMENTS

RESULT 1
C18250 Human placenta cDNA (TFujiwara) Homo sapiens cDNA clone
DEFINITION GEN-559E06 5', mRNA sequence.

ACCESSION C18250
VERSION C18250.1 GI:1579852
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Fujiwara, T., Hirano, H., Hishigaki, H., Horie, M., Kawai, A., Kuga, Y., Kyushiki, H., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Suzuki, M., Takachi, A., Takeda, S., Watanabe, T., Maekawa, H., Nakamura, Y. and Takahashi, E.
TITLE Otsuka cDNA project
JOURNAL Unpublished (1996)
COMMENT Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawachi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.

FEATURES
source
Location/Qualifiers
1..377
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human placenta cDNA (TFujiwara)"
/tissue_type="placenta"

BASE COUNT
91 a 113 c 101 g 72 t
ORIGIN

Query Match 76.6%; Score 81.2; DB 80; Length 377;
Best Local Similarity 90.7%; Pred. No. 2,9e-14;
Matches 98; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

OY 1 aacggcagcgccttcctgcgcgcacccatag--aagccatgcgcgcgcacacacgctca 58
|||||
Db 342 AACGCGACGGCTTCTGCTGCGACCCAAATGCAAGCAATGCGCGACGACGAGCTCA 283
|||||

OY 59 cgcgcgaaggacagcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 106
|||||
Db 282 CGCAGGAAGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 235
|||||

RESULT 2
AA980445 744 bp mRNA EST 27-MAY-1998
ua41d02.f1 Soares mammary-gland_NBMNG Mus musculus cDNA clone
IMAGE:1349283 5' similar to gb:M15169 BETA-2 ADRENERGIC RECEPTOR (HUMAN); mRNA sequence.
AA980445.1 GI:3158981
EST.
ORGANISM Mus musculus
SOURCE house mouse.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 744)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theisinger, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMT Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Apr 7, 1998 this sequence version replaced gi:3035661.
Contact: Maria M/Mouse EST Project
WashU-HMT Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

KEYWORDS	EST.	house mouse.
SOURCE		Mus musculus
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE		1 (bases 1 to 255)
AUTHORS		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Weising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE		The WashU-HMI Mouse EST Project
JOURNAL		Unpublished (1996)
COMMENT		Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:396881 Seq primer: -28ml3 rev2.87 from Amersham High quality sequence stop: 246.
FEATURES		Location/Qualifiers
SOURCE		1..255
		/organism="Mus musculus"
		/strain="C57BL/6J"
		/db_xref="taxon:10090"
		/clone_image:644889"
		/clone_lib="Soares mouse lymph node NbMLN"
		/sex="male"
		/tissue_type="lymph node"
		/dev_stage="4 weeks"
		/lab_host="DH10B"
		/note="Organ: lymph node; Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5
		15]TGTTACCAATCTGAAGTGGAGCGCCGCGATACTTTTTTTTTTTTTTTTTTTT
		3'; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT		61 a 68 c 72 g 54 t
ORIGIN		
		Query Match 30.9%; Score 32.8; DB 22; Length 255;
		Best Local Similarity 64.5%; Pred. No. 4.2;
		Matches 49; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
QY	27	caatagaagccatgcccgcagaccacgacgtcaccgacgaagaagcagagtgtgagtgt 86
Db	181	CCATGAGGAGCCAGGCGCAGACGCCCAAGGAGGAAGCAGCCCGGTGGGGGTGGGTGG 122
QY	87	gggcgtggtgcatcgtc 102
Db	121	GGAGTGGGGCAAGGTC 106
RESULT	5	
LOCUS	A1661315	286 bp mRNA EST 10-MAY-1999
DEFINITION	val5f07.x1 Soares mouse lymph node NbMLN Mus musculus cDNA clone	
ACCESSION	A1661315	
VERSION	A1661315.1	GI:4764898
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Streptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.			
TITLE	The Mashu-NCI Mouse EST Project 1999			
JOURNAL	Unpublished (1999)			
COMMENT	On Jun 22, 1998 this sequence version replaced gi:3247027. Other-ESTs: val5f07.y1 Contact: Marra M/washu-NCI Mouse EST Project 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lln.gov) for further information. MG1:448509 This clone was previously sequenced on the 5' end only, this new data is from the 3' end. Location/Qualifiers 1. 286 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:723013" /clone_1ib="Soares mouse lymph node NBDLN" /sex="male" /tissue.type="lymph node" /dev_stage="4 weeks" /lab_host="DH10B" /note="Organ: lymph node; Vector: pT7T3D-Pac (Pharmacia) and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertando Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."			
BASE COUNT	66 a	82 c	78 g	60 t
ORIGIN	TGTTACCAATCTGAAAGTGGGAGCGCGCGATCTTTTCTTTTTTTTTTTTTTTT 3') double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertando Jordan. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."			
Query Match	30.9%	Score 32.8;	DB 42;	Length 286;
Best Local Similarity	64.5%;	Pred. No. 4.2;		
Matches	49;	Conservative 0;	Mismatches 27;	Indels 0;
Gaps	0;			
Oy	27	caatgaagccatgctgcgcgcagacagctcagcagcaaaaggcaggggtgtgtgtgt 86		
Ddb	68	CCATGAGAGCCCAAGGCTCCAGACCCCAAGGAGCAAGCCCGCTCGGGGtgggtggg 127		
Oy	87	ggcagtgagcatcgtc 102		
Ddb	128	GGAGTGGGGCAAGGTC 143		
RESULT	6			
AA470234/c				
LOCUS	AA470234	370 bp	mRNA	EST
DEFINITION	v090h08.r1 Soares mouse NBMH Mus musculus cDNA clone IMAGE:807903			
ACCESSION	AA470234			
VERSION	AA470234.1 GI:2197543			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			

[illegible]

```

SOURCE
1
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="6430513A21"
/clone_1lb="RIKEN full-length enriched, adult male
olfactory brain"
/sex="male"
/tissue_type="olfactory brain"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGGAGGAGGAGATCCAGAGGCTCTTTTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse

```

BASE COUNT	63 a	58 c	50 g	52 t
ORIGIN				

Oy 27 *caatgaaagccatctggcgcgcgcgcacgaagctcagcaagaagaagagacgagctctggatgt* 86
 Db 156 *ccattgagaccctcaatgctcccaaaaccccaaggacagcaccgccctcgctggcggtggcggtggg* 97
 Oy 87 *gggcattgggcattcgtc* 102
 Db 96 *gcattggcgcccaagcttc* 81

ACCESSION	AI838527
VERSION	AI838527.1
	GI:5472705

AUTHORS	TITLE
BONALDO, M.F., LEMMO	Normalization and S

Contact: Chin, H
National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA

Tel: 301 443 1706
Fax: 301 443 9890

Fax: 301 443 9890
Email: MEST@n1h.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the non-normalized prefrontal cortex library cDNA Library Preparation:

M.B. Scores Lab Clone distribution: NIH BMAP CDNA clones will be made available by the means that is soon to be determined. When NIH

determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.

Seq primer: M13 Forward
POLYA-Yes.

```

FEATURES
source      Location/Qualifiers
1. .305

```

```

/organism="Mus musculus"
/strain="C57BL/6J"

```

```
/db_xref="taxon:10090"
/clone="UI-M-AL0-abp-a-12-0-UI"
```

```
/clone_1lb-NIH_BMAP_MCO
/dev_stage="27-32 days"
```

```
/lab_host="DH10B (Life Technologies)"
```

[illegible]

Db 305 GCCAGGTGCCAGACCCCCCAAGGAGCAGCCCCGGTCGGGGGTGGGGGTGGGGAGTGGG 246

95 gcatcgtc 102

Db 245 GCAAGGTC 238

RESULT	15
AA607055/c	
LOCUS	AA607055 354 bp mRNA EST 30-SEP-1997
DEFINITION	v93564.r.1 Knowles Solter mouse blastocyst B1 Mus musculus CDNN
clone IMAGE:1006014 5', mRNA sequence.	

ACCESSION	AA607055
VERSION	AA607055.1
	GI:2455948

REIMONDS	EST.
SOURCE	house mouse.

mus musculus
Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa;

REFERENCE
1 (bases 1 to 354)
mamillaria; eulheria;
mouderia; scotlogiahl; mulure; muline; mus.

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Maida, M., Miller, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, J.,
Aronson

Scheiwe, R., Stepien, M., Jan, F., Underwood, N., Moore, B., Thetis, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

WALSH, R.
The Washu-HMI Mouse EST Project

On Jan 14, 1998 this sequence replaced g1:1877600
unpublished (1998)
JOURNAL COMMENT

contact: Maria M/mouse Est Project
washu-HHMI Mouse EST Project

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 288 1800
Fax: 314 286 1810

This clone is available royalty-free through LNL; contact the email: mousses@watsol.wustl.edu

IMAGE CONSOLIDUM (image.gov) for further information.
MGI:570230.

```

FEATURES
source
1. .354

```

/organism= mus musculus
/strain="B6D2 F1/J"

```
/add_xiel=caxon:1000
/clone="IMAGE:10060"
```

/tissue_type="blast

```
/dev_stage= emuldyo
/lab_host="DH10B"
```

```

/move-organ: embryo
Site_2: Salt; Clone

```

11011 000 Diastocys
5'-CGTCGACCGTCGACC

CLONED INTO ONE NOU

Wed Sep 13 11:13:55 2000

us-09-542-718-1.rst

Page 10

Technologies), Two different size selections: B1 (larger inserts) and B3."	
BASE COUNT	78 a 95 c 101 g 80 t
ORIGIN	

Query Match	27.9%	Score 29.6;	DB 28;	Length 354;
Best Local Similarity	61.8%	Pred. No. 37;		
Matches 47;	Conservative	0;	Mismatches 29;	Indels 0;
				Gaps 0;

[illegible]

```

OY      87  ggcatggygcacgctc 102
          ||      |||||      |||
Db      236 GGAGTGGGGCAGAGTC  221

```

Search completed: September 12, 2000, 22:46:38
Job time: 3780 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:02:38 : Search time 930.45 seconds
(without alignments)
36.444 Million cell updates/sec

Title: US-09-542-718-2

Perfect score: 19

Sequence: 1 aacggcagcagcctcttcgc 19

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_om:*
6: gb_om:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*
41: gb_ph:*
42: gb_ph:*
43: gb_ph:*

44: gb_hc97:*
45: em_hc91:*
46: em_hc92:*
47: em_hc93:*
48: em_hc94:*
49: gb_ph:*
50: gb_ph:*
51: gb_ph:*
52: gb_ph:*
53: gb_ph:*
54: gb_ph:*
55: gb_ph:*
56: gb_ph:*
57: gb_ph:*
58: gb_ph:*
59: gb_ph:*
60: gb_ph:*
61: gb_ph:*
62: em_hc94:*
63: em_hc95:*
64: em_hc96:*
65: em_hc97:*
66: em_hc98:*
67: gb_hc918:*
68: gb_hc919:*
69: gb_hc920:*
70: gb_hc921:*
71: gb_hc922:*
72: gb_hc923:*
73: gb_hc924:*
74: gb_hc925:*
75: gb_hc926:*
76: gb_hc927:*
77: gb_hc928:*
78: gb_hc929:*
79: gb_hc930:*
80: gb_hc931:*
81: gb_v11:*
82: gb_v12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	19	100.0	1242	11	AF022853 Homo sapi
2	19	100.0	1242	11	AF022854 Homo sapi
3	19	100.0	1242	11	AF022855 Homo sapi
4	19	100.0	1242	11	AF022856 Homo sapi
5	19	100.0	1286	39	AF020305 Homo sapi
6	19	100.0	1290	39	AF169225 Homo sapi
7	19	100.0	1290	39	AF203386 Homo sapi
8	19	100.0	1320	11	MACB2AR
9	19	100.0	1970	9	HSBAR
10	19	100.0	2305	9	HSBAR
11	19	100.0	3451	10	HUMADBRRA
12	19	100.0	3458	10	HUMADBRRA
13	19	100.0	132858	43	AC011354
14	19	100.0	134419	43	AC011354
15	16.4	86.3	720	8	CNS01900
16	16.4	86.3	4146	2	CNS01C22
17	16.4	86.3	6703	1	AF057063
18	16.4	86.3	6703	2	AF057063
19	16.4	86.3	13567	2	AF031898
20	16.4	86.3	37578	33	LMFL7171
21	16.4	86.3	76728	39	AC027185
22	16.4	86.3	108932	69	AF067845
23	16	84.2	660	8	CNS01BHL
24	16	84.2	191281	69	AC022441

25 15.8 83.2 261 12 AF053951 AF053951 Mus muscu
 26 15.8 83.2 287 12 AF020681 AF020681 Mus muscu
 27 15.8 83.2 459 12 AF053948 AF053948 Mus muscu
 28 15.8 83.2 699 12 AF053954 AF053954 Mus muscu
 29 15.8 83.2 1298 12 CEU73206 CEU73206 Mus muscu
 30 15.8 83.2 1437 2 AF170343 AF170343 Mus muscu
 31 15.8 83.2 1437 2 AF165984 AF165984 Mus muscu
 32 15.8 83.2 1536 3 AF192345 AF192345 Mus muscu
 33 15.8 83.2 1791 12 AF010284 AF010284 Mus muscu
 34 15.8 83.2 1868 12 AF134836 AF134836 Mus muscu
 35 15.8 83.2 1948 3 CFB2AR AF005936 AF005936 Mus muscu
 36 15.8 83.2 2315 12 AF059336 AF059336 Mus muscu
 37 15.8 83.2 2679 5 AF153438 AF153438 Rhododact
 38 15.8 83.2 2679 5 A65720 A65720 Sequence 1
 39 15.8 83.2 3035 4 XLU95094 U95094 Xenopus lae
 40 15.8 83.2 4928 12 MMB2ARG X15643 Mouse gene
 41 15.8 83.2 5288 3 PIGB2AR AF000134 Sus scrof
 42 15.8 83.2 12554 2 AE000106 AE000106 Rhizodiu
 43 15.8 83.2 13907 12 AB013129 AB013129 Mus muscu
 44 15.8 83.2 43632 1 SC6A5 AL049485 Streptomy
 45 15.8 83.2 63644 43 AC021508 AC021508 Homo sapi

ALIGNMENTS

RESULT 1
 LOCUS AF022953 1242 bp DNA PRI 30-OCT-1997
 DEFINITION Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
 ACCESSION AF022953
 VERSION AF022953.1 GI:2570526
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Reihaus, E., Innis, M., MacIntyre, N. and Liggett, S.B.
 TITLE Mutations in the gene encoding for the beta 2-adrenergic receptor
 in normal and asthmatic subjects
 JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
 MEDLINE 93192047
 REFERENCE
 AUTHORS Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
 TITLE A polymorphism of the human beta 2-adrenergic receptor within the
 fourth transmembrane domain alters ligand binding and functional
 properties of the receptor
 JOURNAL J. Biol. Chem. 268 (31), 23116-23121 (1993)
 MEDLINE 94043092
 REFERENCE
 AUTHORS Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
 TITLE Amino-terminal polymorphisms of the human beta 2-adrenergic
 receptor impart distinct agonist-promoted regulatory properties
 Biochemistry 33 (32), 9414-9419 (1994)
 JOURNAL 94347707
 MEDLINE 94347707
 REFERENCE
 AUTHORS Liggett, S.B. and Green, S.A.
 TITLE Direct Submission
 JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
 Ave M670564, Cincinnati, OH 45267-0564, USA
 FEATURES
 source
 1. 1242
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. 1242
 /gene="ADRB2"
 1. 1242
 /gene="ADRB2"
 /codon_start=1
 /product="beta2-adrenergic receptor"

/protein_id="AA82148.1"
 /db_xref="GI:2570527"
 /translation="MGPGNGSAFLAPNCSHAPDHVYQVDEYVVGMYMSLIV
 LAIVGVAVITVALAKPERLQVNTFTSLACADLVWGLAVYPERAAHIIMKMTFG
 NFECEFWTSDIVLAVASIEITLCVAVDRFAATSPKQSLTKKARVITLWVAV
 SGLSFLPIQHWYRATHOBAINCYNAMETCDDEFTNQAVAIASIVSFVPLVWV
 YSRPQAKRQIQIDKSEGRHVONLSQVQGRGCHGRSSKCLKEHNAKLTLG
 IIMGTFTLCMLPFTIVNIVHVIDNLRKRVYILLWIGYVNSGFNPLVYCRSPDFRI
 AFQELICLRSSLAAGNGISSNNTGDSGVIHQEKEKKLLCEDLPGETDEYGHQG
 TVPSDNDISQGRNCSTNDLSL"
 46
 /gene="ADRB2"
 /note="Arg16 to Gly polymorphism"
 /replace="a"
 variation
 BASE COUNT 275 a 331 c 326 g 310 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 11; Length 1242;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
 LOCUS AF022954 1242 bp DNA PRI 30-OCT-1997
 DEFINITION Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
 ACCESSION AF022954
 VERSION AF022954.1 GI:2570528
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Reihaus, E., Innis, M., MacIntyre, N. and Liggett, S.B.
 TITLE Mutations in the gene encoding for the beta 2-adrenergic receptor
 in normal and asthmatic subjects
 JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
 MEDLINE 93192047
 REFERENCE
 AUTHORS Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
 TITLE A polymorphism of the human beta 2-adrenergic receptor within the
 fourth transmembrane domain alters ligand binding and functional
 properties of the receptor
 JOURNAL J. Biol. Chem. 268 (31), 23116-23121 (1993)
 MEDLINE 94043092
 REFERENCE
 AUTHORS Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
 TITLE Amino-terminal polymorphisms of the human beta 2-adrenergic
 receptor impart distinct agonist-promoted regulatory properties
 Biochemistry 33 (32), 9414-9419 (1994)
 JOURNAL 94347707
 MEDLINE 94347707
 REFERENCE
 AUTHORS Liggett, S.B. and Green, S.A.
 TITLE Direct Submission
 JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
 Ave M670564, Cincinnati, OH 45267-0564, USA
 FEATURES
 source
 1. 1242
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. 1242
 /gene="ADRB2"
 1. 1242
 /gene="ADRB2"
 /codon_start=1

```

/product="beta2-adrenergic receptor"
/protein_id="AAB82149.1"
/db_xref="GI:2570529"
/translation="MGPGNGSAFLAPNRSHAPDHYTOERDEYVWVGIMVSLIV
LAIVGNLVITAIKFERLOVTNYFTSLACADLVGMLAVPFGAAHIIKMMTFG
NMCSEFMTSIDVCTASIERLCVAVDRYFATSPFYQSILTKNKAHVIIAMWIV
SGTSEFLPIQHMWYRATHOEAINCYANETCCDFPTNOAYAIASSIVSPVPIVWV
YSRVEAKRQLOKIDKSGRPHVONLSQVEDGRTGHLRRSKFCLKKEHAKLTIG
IIMGFTLCMLPEFTIVNIVHYIDNLIRKEYIILLNWIGYVNSGFNPLIYCRSPRI
AFQELCLRRSLKAVNGYSSNGNTGSGYHVEQEKLLCEDLPCTEDFVGHOG
TVPSNDISQGRNCSTNDLSL"
79
/gene="ADRB2"
/note="Gln27 to Glu polymorphism"
/replace="c"
BASE COUNT      276 a      330 c      326 g      310 t
ORIGIN
Query Match      100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. NO. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 aacggcagcgccctcttgc 19
        |||||||
Db      16 AACGCGACGCCCTCTTGC 34

RESULT 3
AF022955      1242 bp      DNA      PRI      30-OCT-1997
LOCUS      Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION
ACCESSION      AF022955
VERSION      AF022955.1 GI:2570530
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS      1 (bases 1 to 1242)
      Relmsaus,E., Innis,M., MacIntyre,N. and Liggett,S.B.
      Mutations in the gene encoding for the beta 2-adrenergic receptor
      in normal and asthmatic subjects
      Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL
MEDLINE      93192047
AUTHORS      2 (bases 1 to 1242)
      Green,S.A., Cole,G., Jacinto,M., Innis,M. and Liggett,S.B.
      A polymorphism of the human beta 2-adrenergic receptor within the
      fourth transmembrane domain alters ligand binding and functional
      properties of the receptor
      J. Biol. Chem. 268 (31), 23116-23121 (1993).
JOURNAL
MEDLINE      94043092
AUTHORS      3 (bases 1 to 1242)
      Green,S.A., Turki,J., Innis,M. and Liggett,S.B.
      Amino-terminal polymorphisms of the human beta 2-adrenergic
      receptor impart distinct agonist-promoted regulatory properties
      Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL
MEDLINE      94347707
AUTHORS      Erratum:[published erratum appears in Biochemistry 1994 Nov
      29;33(47):14368]]
      4 (bases 1 to 1242)
      Liggett,S.B. and Green,S.A.
      Direct Submission
      Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
      Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
source
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADRB2"
1..1242
/gene="ADRB2"
CDS

```

```

/codon_start=1
/product="beta2-adrenergic receptor"
/protein_id="AAB82150.1"
/db_xref="GI:2570531"
/translation="MGPGNGSAFLAPNRSHAPDHYTOERDEYVWVGIMVSLIV
LAIVGNLVITAIKFERLOVTNYFTSLACADLVGMLAVPFGAAHIIKMMTFG
NMCSEFMTSIDVCTASIERLCVAVDRYFATSPFYQSILTKNKAHVIIAMWIV
SGTSEFLPIQHMWYRATHOEAINCYANETCCDFPTNOAYAIASSIVSPVPIVWV
YSRVEAKRQLOKIDKSGRPHVONLSQVEDGRTGHLRRSKFCLKKEHAKLTIG
IIMGFTLCMLPEFTIVNIVHYIDNLIRKEYIILLNWIGYVNSGFNPLIYCRSPRI
AFQELCLRRSLKAVNGYSSNGNTGSGYHVEQEKLLCEDLPCTEDFVGHOG
TVPSNDISQGRNCSTNDLSL"
100
/gene="ADRB2"
/note="Val34 to Met polymorphism"
/replace="g"
BASE COUNT      277 a      331 c      324 g      310 t
ORIGIN
Query Match      100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. NO. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 aacggcagcgccctcttgc 19
        |||||||
Db      16 AACGCGACGCCCTCTTGC 34

RESULT 4
AF022956      1242 bp      DNA      PRI      30-OCT-1997
LOCUS      Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION
ACCESSION      AF022956
VERSION      AF022956.1 GI:2570532
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS      1 (bases 1 to 1242)
      Relmsaus,E., Innis,M., MacIntyre,N. and Liggett,S.B.
      Mutations in the gene encoding for the beta 2-adrenergic receptor
      in normal and asthmatic subjects
      Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL
MEDLINE      93192047
AUTHORS      2 (bases 1 to 1242)
      Green,S.A., Cole,G., Jacinto,M., Innis,M. and Liggett,S.B.
      A polymorphism of the human beta 2-adrenergic receptor within the
      fourth transmembrane domain alters ligand binding and functional
      properties of the receptor
      J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL
MEDLINE      94043092
AUTHORS      3 (bases 1 to 1242)
      Green,S.A., Turki,J., Innis,M. and Liggett,S.B.
      Amino-terminal polymorphisms of the human beta 2-adrenergic
      receptor impart distinct agonist-promoted regulatory properties
      Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL
MEDLINE      94347707
AUTHORS      Erratum:[published erratum appears in Biochemistry 1994 Nov
      29;33(47):14368]]
      4 (bases 1 to 1242)
      Liggett,S.B. and Green,S.A.
      Direct Submission
      Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
      Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
source
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADRB2"
1..1242
/gene="ADRB2"
CDS

```

```

/gene="ADRB2"
/codon_start=1
/product="beta2-adrenergic receptor"
/db_xref="GI:2570533"
/translation="MGQPGNGSAFLAPNRSAPHDVYQQRDEWVYVGMGIMSLIV
LAIVGNVLTALIAKFERLQVTNFTSLACADLVMLAVPFGAHIILKMWTEG
NFWCEFWTSIDVLCVTAIEITLCVAVDRFAITSPKQSLTKRKARVITLWVIV
SGLISFLPIOMHWYRATHOEAINCYNANETCCDEFNQAVAIASSIVSFYPLVIMVY
YSRPOEAKROLKIDKISSEGRHVONLSOVEDGRTGHLRSSKFCLEKHAALKTIG
IMGTFTLCMLPFTIVNIYVIONDLIRKRVYILLMWIGYVNSGFNPILYCRSPDRI
AFQELCLRRSSLKATGNGYSSNGNTGBOSGIHVEDEKKNLLCEDLPGETEDFVGHOG
TVPSDNDISQGRNCSTNDLSL"
491
/variation
/gene="ADRB2"
/nc="Thr164 to Ile polymorphism"
/replace="C"
BASE COUNT      276 a      330 c      325 g      311 t
ORIGIN

Query Match      100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 aacggcagcgccctcttcgc 19
        |||
        16 AACGGCAGCGCCTTCTTGC 34

RESULT 5
AF202305      1286 bp      DNA      PRI      14-DEC-1999
DEFINITION    Homo sapiens beta-2 adrenergic receptor gene, complete cds.
ACCESSION     AF202305
VERSION       AF202305.1 GI:6573152
KEYWORDS
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 1286)
AUTHORS      Rupert,J.L. and Hochachka,P.W.
TITLES       Beta-2-adrenergic receptor allele frequencies in two native
              American populations
              Unpublished
              2 (bases 1 to 1286)
              Rupert,J.L. and Hochachka,P.W.
JOURNAL       Submitted (04-NOV-1999) Zoology, University of British Columbia,
AUTHORS      6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
TITLES       Location/Qualifiers
              1. 1286
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="5"
              /map="5q31-q33"
              /note="Isolated from Quechua speaking Native American"
              <35. >1276
              /product="beta-2 adrenergic receptor"
              35. 1276
              /codon_start=1
              /product="beta-2 adrenergic receptor"
              /protein_id="AAF17569.1"
              /db_xref="GI:6573153"
              /translation="MGQPGNGSAFLAPNRSAPHDVYQQRDEWVYVGMGIMSLIV
              LAIVGNVLTALIAKFERLQVTNFTSLACADLVMLAVPFGAHIILKMWTEG
              NFWCEFWTSIDVLCVTAIEITLCVAVDRFAITSPKQSLTKRKARVITLWVIV
              SGLISFLPIOMHWYRATHOEAINCYNANETCCDEFNQAVAIASSIVSFYPLVIMVY
              YSRPOEAKROLKIDKISSEGRHVONLSOVEDGRTGHLRSSKFCLEKHAALKTIG
              IMGTFTLCMLPFTIVNIYVIONDLIRKRVYILLMWIGYVNSGFNPILYCRSPDRI
              AFQELCLRRSSLKATGNGYSSNGNTGBOSGIHVEDEKKNLLCEDLPGETEDFVGHOG
              TVPSDNDISQGRNCSTNDLSL"
CDS
mrna
CDS

```

```

BASE COUNT      282 a      347 c      334 g      320 t      3 others
ORIGIN

Query Match      100.0%; Score 19; DB 39; Length 1286;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 aacggcagcgccctcttcgc 19
        |||
        50 AACGGCAGCGCCTTCTTGC 68

RESULT 6
AF169225      1290 bp      DNA      PRI      10-AUG-1999
DEFINITION    Homo sapiens beta-2-adrenergic receptor gene, complete cds.
ACCESSION     AF169225
VERSION       AF169225.1 GI:5714687
KEYWORDS
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 1290)
AUTHORS      Rupert,J.R. and Hochachka,P.W.
TITLES       Beta-2-adrenergic receptor allele frequencies in two native
              American populations
              Unpublished
              2 (bases 1 to 1290)
              Rupert,J.R. and Hochachka,P.W.
JOURNAL       Submitted (14-JUL-1999) Zoology, University of British Columbia,
AUTHORS      6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
TITLES       Location/Qualifiers
              1. 1290
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="5"
              /map="5q31-q33"
              /cell_type="lymphocyte"
              /tissue_type="Blood"
              /note="Isolated from a Quechua speaking native American
              heterozygous for a known C/T mutation"
              17. 1258
              /codon_start=1
              /product="beta-2-adrenergic receptor"
              /protein_id="AA048036.1"
              /db_xref="GI:5714688"
              /translation="MGQPGNGSAFLAPNRSAPHDVYQQRDEWVYVGMGIMSLIV
              LAIVGNVLTALIAKFERLQVTNFTSLACADLVMLAVPFGAHIILKMWTEG
              NFWCEFWTSIDVLCVTAIEITLCVAVDRFAITSPKQSLTKRKARVITLWVIV
              SGLISFLPIOMHWYRATHOEAINCYNANETCCDEFNQAVAIASSIVSFYPLVIMVY
              YSRPOEAKROLKIDKISSEGRHVONLSOVEDGRTGHLRSSKFCLEKHAALKTIG
              IMGTFTLCMLPFTIVNIYVIONDLIRKRVYILLMWIGYVNSGFNPILYCRSPDRI
              AFQELCLRRSSLKATGNGYSSNGNTGBOSGIHVEDEKKNLLCEDLPGETEDFVGHOG
              TVPSDNDISQGRNCSTNDLSL"
491
/variation
/replace="C/T"
BASE COUNT      287 a      349 c      331 g      322 t      1 others
ORIGIN

Query Match      100.0%; Score 19; DB 39; Length 1290;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 aacggcagcgccctcttcgc 19
        |||
        32 AACGGCAGCGCCTTCTTGC 50

RESULT 7

```

AF203386 LOCUS AF203386 1290 bp DNA PRI 28-DEC-1999
 DEFINITION Homo sapiens beta-2 adrenergic receptor (ADRB2) gene, complete cds.
 ACCESSION AF203386
 VERSION AF203386.1 GI:6636495
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1290)
 AUTHORS Rupert, J.L., and Hochachka, P.W.
 TITLE Beta-2 adrenergic receptor allele frequencies in two Native American populations
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1290)
 AUTHORS Rupert, J.L., and Hochachka, P.W.
 TITLE Direct Submission
 JOURNAL Submitted (09-NOV-1999) Zoology, University of British Columbia, 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
 FEATURES
 source
 1..1290
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q32-q34"
 /sex="female"
 /note="Isolated from a Quechuan-speaking Native American"
 26..1267
 /gene="ADRB2"
 26..1267
 /gene="ADRB2"
 /note="catecholamine receptor"
 /codon_start=1
 /product="beta-2 adrenergic receptor"
 /protein_id="AAF20199.1"
 /db_xref="GI:6636496"
 /translation="MGPGNGSAPFLAPNGSHAPDHYTOERDEAVVVGMIYMSLI
 LAIVGNLVITAIAKPERLQTVTVETISLACADLVGAVPFGAAHILKMTFG
 NWCPEWTSIDVLCVTASLETICVAVDVFYITSPFKQSLITNKARVITLWMIV
 SGLTSPFLQHMWRATIDQEAINCYNANECDFPTNOAVAIASSIVSPVPIVMIV
 YSRVDEAKRQLOKIDKSEGRHVNLSQVEDDGRTHGLRSSKFCLEHKAALTLG
 IIMGFTLCMLPEFTIVNIVYIODNLIKEVYILLNWGVYVSGNPFLYCSPPERI
 AFOELLCILRRSSILKACNGYSNSNGTGEOSGYHVEQEKNKLLCEDLPETEDFVGHG
 TVPSDNISQGCNSGTNDLSL"
 BASE COUNT 288 a 345 c 333 g 324 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 39; Length 1290;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 aacggcagcgactctctgc 19
 ||||||||||||||||
 Db 41 AACGCGACGCCCTTCTTGC 59
 RESULT 8
 LOCUS MACB2AR 1320 bp mRNA PRI 02-OCT-1995
 DEFINITION Macaca mulatta beta-2 adrenergic receptor (B2AR) mRNA, complete
 cds
 ACCESSION L38905
 VERSION L38905.1 GI:1004338
 KEYWORDS beta-2 adrenergic receptor.
 SOURCE Macaca mulatta cDNA to mRNA.
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
 Macaca.
 REFERENCE 1 (bases 1 to 1320)
 AUTHORS Amend, A.M. and Guan, X.M.

TITLE Cloning, sequencing, and expression of the rhesus monkey beta 2
 JOURNAL adrenergic receptor
 MEDLINE DNA Cell Biol. 14 (9), 753-757 (1995)
 FEATURES 95398843
 source
 1..1320
 /organism="Macaca mulatta"
 /db_xref="taxon:9544"
 /note="genomic DNA for this receptor sequenced and found to be identical with the cDNA"
 1..40
 /gene="B2AR"
 1..1320
 /gene="B2AR"
 1..1320
 /gene="B2AR"
 41..1288
 /gene="B2AR"
 /codon_start=1
 /product="beta-2 adrenergic receptor"
 /protein_id="AAC41914.1"
 /db_xref="GI:1004339"
 /translation="MGPGNGSAPFLAPNGSHAPDHYTOERDEAVVVGMIYMSLI
 LAIVGNLVITAIAKPERLQTVTVETISLACADLVGAVPFGAAHILKMTFG
 NWCPEWTSIDVLCVTASLETICVAVDVFYITSPFKQSLITNKARVITLWMIV
 SGLTSPFLQHMWRATIDQEAINCYNANECDFPTNOAVAIASSIVSPVPIVMIV
 YSRVDEAKRQLOKIDKSEGRHVNLSQVEDDGRTHGLRSSKFCLEHKAALTLG
 IIMGFTLCMLPEFTIVNIVYIODNLIKEVYILLNWGVYVSGNPFLYCSPPERI
 AFOELLCILRRSSILKACNGYSNSNGTGEOSGYHVEQEKNKLLCEDLPETEDFVGH
 OCTVPSDNISQGCNSGTNDLSL"
 1289..1320
 /gene="B2AR"
 BASE COUNT 289 a 365 c 343 g 323 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 11; Length 1320;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 aacggcagcgactctctgc 19
 ||||||||||||||||
 Db 56 AACGCGACGCCCTTCTTGC 74
 RESULT 9
 LOCUS HSBAR 1970 bp mRNA PRI 12-SEP-1993
 DEFINITION Human mRNA for brain beta-adrenergic receptor.
 ACCESSION X04827
 VERSION X04827.1 GI:29372
 KEYWORDS beta-adrenergic receptor.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1970)
 AUTHORS Chung, F.Z., Lentjes, K.U., Gocayne, J., Fitzgerald, M., Robinson, D.,
 Kerlavage, A.R., Fraser, C.M. and Venter, J.C.
 TITLE Cloning and sequence analysis of the human brain beta-adrenergic
 receptor. Evolutionary relationship to rodent and avian
 beta-receptors and porcine muscarinic receptors
 JOURNAL FEBS Lett. 211 (2), 200-206 (1987)
 MEDLINE 87105974
 REFERENCE 2 (bases 1 to 1970)
 AUTHORS Kerlavage, A.R.
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-1987) to the EMBL/GenBank/DDay databases
 COMMENT Substantial corrections are reported in [2]
 DATA Kindly reviewed (22-SEP-1987) by Kerlavage A.R.
 FEATURES
 source
 1..1970
 /organism="Homo sapiens"

```

CDS
    /db_xref="taxon:9606"
    /clone_lib="neonatal human brain stem"
    178..1419
    /note="beta-adrenergic receptor (AA 1-413)"
    /codon_start=1
    /protein_id="CAA28511.1"
    /db_xref="GI:29373"
    /db_xref="SWISS-PROT:P07550"
    /translation="MGOPNGSAFLAPRSHAPDHDTQERDEWVWVGIMVLSLIV
    LAIVGNVITAIKAFERLOTVTNYFTSLACADIVMGIAVVPFGAAHLLMKMTFG
    NFWCEWTSIDVLCVATSIETLCVIAVDREYFAITSPKQSLTKKARVITLIMVIV
    SGLTSFLPIOMHYRATHOEA INCYANECDFPNOAYAIASSIVSEVPLIVMVF
    YSRVDEAKROLKIDKSEGRFVQNLSEODGCRGHGRSSKCLKHKALKITG
    IIMGFTLCMLPEFTIVNIHVIODNLRKQVITLIMIGVNSGFNPLIYCRSPDRI
    AFQELCLRRSLKAYGNGYSNGNTGEOSGYHVEQEKENKLLCEDLPETDEVFVHGQ
    TVPSDNIDSGRNCSTNDSL"
    misc_feature
        /note="pot. glucocorticoid-responsive element"
        794..799
        /note="pot. glucocorticoid-responsive element"
        965..970
        /note="pot. glucocorticoid-responsive element"
        1459..1464
        /note="pot. glucocorticoid-responsive element"
        1491..1496
        /note="pot. polyA signal"
        1502..1507
        /note="pot. polyA signal"
        1952..1957
        /note="pot. polyA signal"
        1970
        /note="polyA site"
        508 c 482 g 521 t
BASE COUNT 459 a 508 c 482 g 521 t
ORIGIN
Query Match 100.0%; Score 19; DB 9; Length 1970;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 aacggcagcgctctcttcg 19
    |||||||
Db 193 AACGGCAGCGCTCTCTGC 211

RESULT 10
LOCUS HSRAR 2305 bp DNA PRI 12-SEP-1993
DEFINITION Human gene for beta-adrenergic receptor (beta-2 subtype).
ACCESSION Y00106
VERSION Y00106.1 GI:29370
KEYWORDS beta-adrenergic receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2305)
AUTHORS Schofield,P.R., Rhee,L.M. and Peralta,E.G.
TITLE Primary structure of the human beta-adrenergic receptor gene
JOURN. Nucleic Acids Res. 15 (8), 3636 (1987)
MEDLINE 87203400
REFERENCE 2 (bases 1 to 2305)
AUTHORS Schofield,P.R.
TITLE Direct Submission
JOURN. Submitted (20-OCT-1987) to the EMBL/GenBank/DBJ databases
FEATURES
    location/Qualifiers
    1..2305
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="Manatis human"
    /clone="lambda/betaAR17"
    794..2035
    /note="beta-adrenergic receptor (AA 1 - 413)"
    /codon_start=1
    /protein_id="CAA68289.1"
CDS

```

```

    /db_xref="GI:29371"
    /db_xref="SWISS-PROT:P07550"
    /translation="MGOPNGSAFLAPRSHAPDHDTQERDEWVWVGIMVLSLIV
    LAIVGNVITAIKAFERLOTVTNYFTSLACADIVMGIAVVPFGAAHLLMKMTFG
    NFWCEWTSIDVLCVATSIETLCVIAVDREYFAITSPKQSLTKKARVITLIMVIV
    SGLTSFLPIOMHYRATHOEA INCYANECDFPNOAYAIASSIVSEVPLIVMVF
    YSRVDEAKROLKIDKSEGRFVQNLSEODGCRGHGRSSKCLKHKALKITG
    IIMGFTLCMLPEFTIVNIHVIODNLRKQVITLIMIGVNSGFNPLIYCRSPDRI
    AFQELCLRRSLKAYGNGYSNGNTGEOSGYHVEQEKENKLLCEDLPETDEVFVHGQ
    TVPSDNIDSGRNCSTNDSL"
    misc_feature
        /note="N-linked glycosylation site"
        809..817
        /note="N-linked glycosylation site"
        836..844
        /note="N-linked glycosylation site"
        896..967
        /note="membrane spanning domain I"
        1007..1078
        /note="membrane spanning domain II"
        1114..1180
        /note="membrane spanning domain III"
        1247..1315
        /note="membrane spanning domain IV"
        1385..1450
        /note="membrane spanning domain V"
        1616..1687
        /note="membrane spanning domain VI"
        1712..1774
        /note="membrane spanning domain VII"
        616 c 649 g 545 t
BASE COUNT 495 a 616 c 649 g 545 t
ORIGIN
Query Match 100.0%; Score 19; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 aacggcagcgctctcttcg 19
    |||||||
Db 809 AACGGCAGCGCTCTCTGC 827

RESULT 11
LOCUS HUMADRB 3451 bp mRNA PRI 13-FEB-1996
DEFINITION Human beta-2-adrenergic receptor mRNA, complete cds.
ACCESSION M15169 J02728 M16106
VERSION M15169.1 GI:178201
KEYWORDS adrenergic receptor.
SOURCE Homo sapiens (clone: pTF.) (tissue library: Evan Sadler) Placenta
cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3451)
AUTHORS Koblika,B.K., Dixon,R.A., Frielle,T., Dohlman,H.G., Bolanowski,M.A.,
Dixon,R.A., Keller,P., Caron,M.G. and Lefkowitz,R.J.
TITLE Delineation of the intronless nature of the genes for the human and
hamster beta 2-adrenergic receptor and their putative promoter
regions
JOURN. J. Biol. Chem. 262 (15), 7321-7327 (1987)
MEDLINE 87222338
REFERENCE 2 (bases 1399 to 1985)
AUTHORS Koblika,B.K., Dixon,R.A., Frielle,T., Dohlman,H.G.,
Bolanowski,M.A., Sigal,I.S., Yang-Peng,T.L., Francke,U., Caron,M.G.
and Lefkowitz,R.J.
TITLE cDNA for the human beta 2-adrenergic receptor: a protein with
multiple membrane-spanning domains and encoded by a gene whose
chromosomal location is shared with that of the receptor for
platelet-derived growth factor
JOURN. Proc. Natl. Acad. Sci. U.S.A. 84 (1), 46-50 (1987)
MEDLINE 87092393
FEATURES
    location/Qualifiers
    1..3451

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	100.0%;	100.08;	Pred. No. 54;	3451;
Db 1603 AACGCGACGCCCTTCTGC 1621	1 aacgcagcgacctcttcgc 19			
RESULT 12				
HUMADRBRA	LOCUS			
DEFINITION	HUMAN beta-2-adrenergic receptor gene, complete cds.			
ACCESSION	J02960			
VERSION	J02960.1			
KEYWORDS	GI:178203			
SOURCE	adrennergic receptor; beta-2 adrenergic receptor.			
ORGANISM	Homo sapiens (clone: H-beta-R-[9,10,11]) epidermis DNA.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 3458)			
JOURNAL	Emorlie, L.J., Merullo, S., Delavier-Klutchko, C., Kaveri, S.V.,			
	DuRen-Trautmann, O. and Strosberg, A.D.			
	Structure of the gene for human beta-2-adrenergic receptor:			
	expression and promoter characterization			
	Proc. Natl. Acad. Sci. U.S.A. 84 (20), 6995-6999 (1987)			

MEDLINE
 88041037
COMMENT
 Draft entry and computer-readable copy of sequence [1] kindly
 provided by L.J. Emorine, 25-AUG-1987.
FEATURES
 source
 1..3458
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="H-beta-R-[9,10,11]"
 /cell_line="A831"
 /tissue_type="epidermis"
 /map="5q31-q32"
 277..1032
 /note="ORF: putative"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AA88016.1"
 /db_xref="GI:560762"
 /translacion="MFEREYTGLPVCEGSIISARVROYRSTOMETSVSLMPPS:
 ORVPCVCHHVEFLVGLASVPSVSGRVSLDRGDPVPRGFCVRAASVHGLGCVS
 SMAYRKSEHVCGGVPPVCACLGHSRFLPNVGCRCALCLETSSRGAOSROYA
 ATPEKAPAGLAKHTTSSFSPLGARVAGQWPAALOGAVGPRGPQEREGEGRGK
 GEELAPSRLPACMKPVVPRHGGSSPKVLCI"
 1045..3057
 /note="beta-2-adrenergic receptor mRNA (alt.)"
 1055..3057
 /note="beta-2-adrenergic receptor mRNA (alt.)"
 1064..3057
 /note="beta-2-adrenergic receptor mRNA (alt.)"
 1264..2505
 /gene="ADRB2"
 1264..2505
 /gene="ADRB2"
 /codon_start=1
 /db_xref="GDB:G00-120-541"
 /product="beta-2 adrenergic receptor"
 /protein_id="AA88017.1"
 /db_xref="GI:178204"
 /translacion="MGCPGNSAFLLPNGSHAPDHDTVOORDEVWVYKCVNSLIV
 LAIVGNVLVITALAKFERLDQTVNFTTSLACDLVGLAVPFGAAHILMKMTG
 NFMEFMTSIDVLTASIEFLCVIADRFALTPSPKOSLIRKNRARIILMWIV
 SGLSPLPIQMHWRATHQEAINCVAIVNETCCDFPNQAVAIASSIVSEYVLIMVIV
 YSRFEQAKROIOLDKSEGFHONLSQVRODRTGHLRRSPCLKEKKALKITG
 IKMTPTLCHLPRFIVINIVHYTONLIRKEVITILNMGIVNSFPNLICRSDDPII
 AFQDLCLRRSLKAVGVSSNGNTGDSQYHVEQERENKLLEDLPTEDFVHGOG
 TVPSDNDISQGRNCTSDSL"
 777 a 890 c 886 g 905 t
BASE COUNT
 1 bp upstream of EcoRI site; chromosome 5q31-q32.
ORIGIN
 100.0%; Score 19; DB 10; Length 3458;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 0y 1 aacgcagccctctctgc 19
 |||||
 |||||
RESULT 13
AC011354
LOCUS AC011354 132858 bp DNA HTG 28-JAN-2000
DEFINITION Homo sapiens chromosome 5 clone CIT-HSPC_354F19, WORKING DRAFT
SEQUENCE 3 ordered pieces.
AC011354
AC011354.1 GI:6013586
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE
 human.
ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
 1 (bases 1 to 132858)
AUTHORS
 DOE Joint Genome Institute.
TITLE
 Sequencing of Human Chromosome 5

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 132858)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 www.jgi.doe.gov.

COMMENT
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1.132858
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CIT-HSPC_354F19"
 location/Qualifiers
 60532 132858: contig of 72327 bp in length.

BASE COUNT 40121 a 26539 c 25722 g 40476 t

ORIGIN

Query Match 100.0%; Score 19; DB 43; Length 132858;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aacggcagcgcctcttcg 19
 |||||||||||||||||||
 Db 127624 AACGGCAGCGCCTTCTTC 127642

RESULT 14
 AC011334 134419 bp DNA HTG 28-JAN-2000
 LOCUS Homo sapiens chromosome 5 clone CIT-HSPC_235N17, WORKING DRAFT
 DEFINITION
 AC011334
 AC011334.1 GI:6013606
 HTG; HTGS_PHASE2; HTGS_DRAFT.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 134419)
 DOE Joint Genome Institute.
 Sequencing of Human Chromosome 5
 Unpublished
 2 (bases 1 to 134419)
 DOE Joint Genome Institute.
 Direct Submission
 Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 www.jgi.doe.gov.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 SUBMITTED
 AUTHORS
 TITLE
 JOURNAL
 SUBMITTED

COMMENT
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 134419: contig of 134419 bp in length.
 location/Qualifiers

FEATURES

source 1.134419
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CIT-HSPC_235N17"
 location/Qualifiers
 40270 a 28042 c 27433 g 38674 t

BASE COUNT 40270 a 28042 c 27433 g 38674 t

ORIGIN

Query Match 100.0%; Score 19; DB 43; Length 134419;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aacggcagcgcctcttcg 19
 |||||||||||||||||||
 Db 54780 AACGGCAGCGCCTTCTTC 54798

RESULT 15
 AC011334 134419 bp mRNA PLN 02-SEP-1999
 LOCUS Botrytis cinerea strain T4 cDNA library under conditions of
 DEFINITION nitrogen deprivation.
 AC011334
 AC011334.1 GI:5826571
 VERSION
 AL111952
 AL111952.1 GI:5826571
 KEYWORDS cDNA library; nitrogen deprivation.
 SOURCE Botryotinia fuckeliana.
 ORGANISM Botryotinia fuckeliana
 Eukaryota; Fungi; Ascomycota; Euascomycetes; Discomycetes;
 Leotiales; Sclerotiniaceae; Botryotinia.
 1 (bases 1 to 720)
 Bilton, F., Lewis, C., Fortin, D., Pradier, J.M. and Brygoo, Y.
 Direct Submission
 Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,
 78026 Versailles, France
 2 (bases 1 to 720)
 Genoscope.
 Direct Submission
 Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage ;
 CP 5706 91057 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)
 The cDNA library to be analyzed within the framework of this
 project was created using a Botrytis cinerea strain which was grown
 under conditions of nitrogen deprivation, which is the normal
 situation for B. cinerea during its development on its host plant.
 The library was produced in an oriented direction, in the pBSII
 vector.

FEATURES
 source
 1.720
 /organism="Botryotinia fuckeliana"
 /strain="T4"
 /db_xref="taxon:40559"
 /note="Genoscope sequence ID : W30B101"
 location/Qualifiers
 185 a 177 c 159 g 198 t 1 others

BASE COUNT 185 a 177 c 159 g 198 t 1 others

ORIGIN

Query Match 86.3%; Score 16.4; DB 8; Length 720;
 Best Local Similarity 94.4%; Pred. No. 9.7e+02;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 aacggcagcgcctcttcg 18
 |||||||||||||||||||
 Db 80 AACGGCAGCGCCTTCTTG 63

Search completed: September 12, 2000, 23:02:47
 Job time: 3959 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:07:08 : Search time 148.16 Seconds
(without alignments)
32.085 Million cell updates/sec

Title: US-09-542-718-2

Perfect score: 19

Sequence: 1 aacgcgcgcgcctcttcgc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_36:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	1999	1 T93250	Beta-2 adrenalin r
2	19	100.0	3451	1 V52614	Human beta-2-adren
3	17.4	91.6	1400	1 T93249	Beta-2 adrenalin r
4	15.8	83.2	322	1 X33329	PEBP2 alpha A poly
5	15.8	83.2	2679	1 V30468	Canine beta-2 adre
6	15.8	83.2	3334	1 V80117	Osif/Cbfa1 polypep
7	15.8	83.2	6178	1 V80182	Nucleic acid seque
8	15.8	83.2	13907	1 X33247	PEBP2 alpha A gene
9	15.8	83.2	110000	1 V30458.4	Continuation (5 of
10	15.8	83.2	110000	1 V30459.4	Continuation (5 of
11	15.4	81.1	1400	1 Q70142	Human cartilage 11
12	15.4	81.1	8625	1 T45143	Flavobacterium car
13	15.4	81.1	8991	1 X13195	Enterococcus faeca
14	15.4	78.9	1417	1 T87401	AD4/AD3LP sequence
15	15.4	78.9	2329	1 T40031	Human presentin-2
16	15.4	78.9	2329	1 V04669	Human presentin-2
17	15.4	78.9	2326	1 T51253	Human AD4 protein
18	15.4	78.9	2276	1 T87426	Full AD4/AD3LP seq
19	15.4	78.9	4004	1 T51260	Human AD4 gene gen
20	14.8	77.9	4888	1 V22134	Rice glutelin-1 pr
21	14.4	75.8	113	1 T39094	Partial P. puncta
22	14.4	75.8	897	1 V37165	DNA sequence used
23	14.4	75.8	1790	1 O66071	Sequence of corn m
24	14.4	75.8	1912	1 T17715	Heat resistant mal
25	14.4	75.8	3713	1 T17715	Heat resistant mal
26	14.4	75.8	3845	1 O47064	tie truncated rece
27	14.4	75.8	3845	1 O47064	Human phosphatidy
28	14.4	75.8	24379	1 T93095	CDNA encoding plat
29	14.4	75.8	24379	1 V25925	Platelet-derived g
30	14.2	74.7	275	1 X41337	Type B human plate
31	14.2	74.7	501	1 T36046	CDNA encoding huma
32	14.2	74.7	555	1 T36046	Enterococcus faeca
33	14.2	74.7	706	1 T41722	Enterococcus faeca
34	14.2	74.7	1049	1 V59558	PEBP2 alpha A gene
35	14.2	74.7	1049	1 V59558	Prion protein bind

34	14.2	74.7	2400	1 T42859	Choline oxidase ge
35	14.2	74.7	2400	1 T75000	Choline oxidase ge
36	14.2	74.7	2454	1 O81966	Human lysosomal me
37	14.2	74.7	3425	1 V99778	Human phosphatidy
38	14.2	74.7	4544	1 N90355	CDNA encoding plat
39	14.2	74.7	4544	1 T34552	Platelet-derived g
40	14.2	74.7	5427	1 Q27447	Type B human plate
41	14.2	74.7	5719	1 N90388	CDNA encoding huma
42	14.2	74.7	8911	1 X13588	Enterococcus faeca
43	14.2	74.7	27360	1 X13101	Enterococcus faeca
44	13.8	72.6	27	1 X33252	PEBP2 alpha A gene
45	13.8	72.6	549	1 V01476	Prion protein bind

ALIGNMENTS

RESULT 1	1	100.0%	Score 19; DB 1; Length 1999;
ID	T93250	standard; cDNA to mRNA; 1999 BP.	
AC	T93250;		
DR	20-APR-1998 (first entry)		
DE	Beta-2 adrenalin receptor subtype coding sequence.		
KW	Beta-2 adrenalin subtype; cyanophindrol; agonist; antagonist;		
KW	asthmatic disease; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	190..1431	
FT		/*tag= a	
PN	WO9735963-A1.		
PD	02-OCT-1997.		
PR	24-MAR-1997; J00982.		
PR	27-MAR-1996; JP-072914.		
PA	(DAN) DAINIPPON PHARM CO LTD.		
PI	Fujii K, Furutani Y, Kawashima H, Nomura A, Yano K;		
DR	WPI: 97-489627/45.		
DR	P-PSDB: W34320.		
PT	Novel beta-2 adrenalin receptor sub-type - useful for screening for		
PT	agonists and antagonists and researching asthmatic diseases		
PS	Disclosure; Page 27-30; 47pp; Japanese.		
CC	This sequence encodes the protein of the invention. The protein of the		
CC	invention is a beta-2 adrenalin receptor subtype with Kd value of		
CC	approximately 75 pM against 125I-cyanophindrol. The protein can be used in		
CC	screening for agonists and antagonists, which are useful in researching		
CC	asthmatic diseases.		
SQ	Sequence 1999 BP; 477 A; 513 C; 485 G; 524 T;		
Query Match	100.0%; Score 19; DB 1; Length 1999;		
Best Local Similarity	100.0%; Pred. NO. 1.2;		
Matches 19; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 aacgcgcgcgcctcttcgc 19		
DB	205 AACGCAGCGCCTCTTGC 223		
RESULT 2	2	100.0%	Score 19; DB 1; Length 1999;
ID	V52614	standard; cDNA; 3451 BP.	
AC	V52614;		
DR	21-DEC-1998 (first entry)		
DE	Human beta-2-adrenergic receptor cDNA.		
KW	Beta-2-adrenergic receptor; human; asthma; beta-agonist;		
KW	polymorphism; ds.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1588..2829	
FT		/*tag= a	
FT	variation	1633	
FT	/*tag= b		
FT	/note= "A to G substitution, results in Arg16		
FT	to Gly amino acid change"		

PN W09839477-A2.
 PD 11-SEP-1998.
 PR 26-FEB-1998; U03908.
 PA (BGHM) BRIGHM & WOMENS HOSPITAL.
 PI Boushey H, Chinchilli VM, Drazen JM, Fish JE, Ford JG,
 PI Martin RJ;
 DR WPI; 98-506372/43.
 DR P-PSDB; W75777.
 PT Diagnosing asthma patients predisposed to adverse beta-agonist
 PT reactions upon regular administration - by identifying patients
 PT homozygous for allele encoding Arg at position 16 of
 PT beta2-adrenergic receptor protein
 PS Disclosure; Page 33-35; 46pp; English.
 CC This CDNA sequence codes for human beta-2-adrenergic receptor (see
 CC W75777) having an arginine residue at position 16. A novel method
 CC for identifying individuals susceptible to adverse responses to
 CC regular administration of beta-agonists comprises: (a) identifying
 CC in a genomic nucleic acid sample from the individual first and
 CC second alleles of the beta 2-adrenergic receptor gene, and (b)
 CC classifying an individual as susceptible if first and second
 CC alleles both encode Arg at residue 16 of the beta 2-adrenergic
 CC receptor protein. Beta 2-adrenergic receptor gene alleles may be
 CC identified by any known method e.g. denaturing gel electrophoresis
 CC or PCR amplification (see also V52615-17). Identification
 CC preferably comprises amplifying a portion of each allele which
 CC includes the sequence encoding residue 16, and optionally also
 CC comprises determining nucleotide sequences of these portions (e.g.
 CC by automated sequence analysis). The invention identifies a known
 CC polymorphism in the beta 2-adrenergic receptor gene as being linked
 CC to adverse responses to regular beta-agonist administration;
 CC position 16 of the encoded protein can be either Arg or Gly, and
 CC individuals homozygous for Arg16 are more susceptible.
 SQ Sequence 3451 Bp; 790 A; 873 C; 895 G; 893 T;

Query Match 100.0%; Score 19; DB 1; Length 3451;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 aacgcagcagcctcttgc 19
 |||||
 Db 1603 AACGCAGCGCCTCTTGC 1621

RESULT 3
 T93249 T93249 standard; cDNA to mRNA; 1400 BP.
 ID T93249;
 AC T93249;
 DT 20-APR-1998 (first entry)
 DE Beta-2 adrenergic receptor subtype coding sequence.
 KW Beta-2 adrenergic receptor; cyanopindrol; agonist; antagonist;
 KW asthmatic disease; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 101..1348
 FT /tag= a
 FT /transl_except= (pos: 338..340, aa: Xaa)
 FT /note= "Xaa-Leu, Ile"
 FT
 FT W09735963-A1.
 PD 02-OCT-1997.
 PR 24-MAR-1997; J00982.
 PR 27-MAR-1996; JP-072914.
 PA (DAIN) DAINIPPON PHARM CO LTD.
 PI Fujii K, Furutani Y, Kawashima H, Nomura A, Yano K;
 DR WPI; 97-489627/45.
 DR P-PSDB; W34319.
 PT Novel beta-2 adrenergic receptor sub-type - useful for screening for
 PT agonists and antagonists and researching asthmatic diseases
 PS Claim 5; Page 23-26; 47pp; Japanese.
 CC This sequence encodes the protein of the invention. The protein of the
 CC invention is a beta-2 adrenergic receptor subtype with Kd value of
 CC approximately 75 pM against 125I-cyanopindrol. The protein can be used in

CC screening for agonists and antagonists, which are useful in researching
 CC asthmatic diseases.
 SQ Sequence 1400 Bp; 304 A; 402 C; 363 G; 328 T;

Query Match 91.6%; Score 17.4; DB 1; Length 1400;
 Best Local Similarity 94.7%; Pred. No. 7.1;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 aacgcagcagcctcttgc 19
 |||||
 Db 116 AACGCAGCGCCTCTTGC 134

RESULT 4
 X33329/c X33329 standard; cDNA to mRNA; 322 BP.
 ID X33329;
 AC X33329;
 DT 30-JUN-1999 (first entry)
 DE PEBP2 alpha A polynucleotide sequence SEQ ID NO:3.
 KW PEBP2 alpha A gene; expression; regulation; bone disease;
 KW osteoporosis; ds.
 OS Mus sp.
 PN W09911787-A1.
 PD 11-MAR-1999.
 PR 02-SEP-1998; J03920.
 PR 08-APR-1998; JP-114135.
 PR 02-SEP-1997; JP-254250.
 PR 15-OCT-1997; JP-299407.
 PA (SUMU) SUMITOMO PHARM CO LTD.
 PI Fujiwara M, Harada H, Katsumata T, Nakatsuka M,
 PI Ogawa S, Tagashira S;
 DR WPI; 99-243621/20.
 PT DNA regulating expression of PEBP2 alpha gene to produce regulator
 PT protein, useful as promoter for prevention or/and treatment of bone
 PT diseases e.g. osteoporosis.
 PS Example 15; Page 75-76; 118pp; Japanese.
 CC The present invention describes DNA which participates in the regulation
 CC of expression of PEBP2 alpha A gene. The DNA produces a regulator
 CC protein with the activity of promoting bone formation and can serve as a
 CC promoter for prevention and treatment of bone diseases including
 CC osteoporosis. The present sequence represents a polynucleotide sequence
 CC from the present invention.
 SQ Sequence 322 Bp; 95 A; 77 C; 74 G; 76 T;

Query Match 83.2%; Score 15.8; DB 1; Length 322;
 Best Local Similarity 89.5%; Pred. No. 37;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 aacgcagcagcctcttgc 19
 |||||
 Db 114 AACGCAGCGCCTCTTGC 96

RESULT 5
 V30468 V30468 standard; cDNA to mRNA; 2679 BP.
 ID V30468;
 AC V30468;
 DT 14-OCT-1998 (first entry)
 DE Canine beta-2 adrenergic receptor coding sequence.
 KW Canine; beta-adrenergic receptor; brown adipose tissue; probe; human;
 KW hybridisation; ligand; ss.
 OS Canis familiaris.
 FH Key Location/Qualifiers
 FT CDS 169..1416
 FT /tag= a
 FT /product= "beta-2 adrenergic receptor"
 FT
 FT W09735973-A2.
 PD 02-OCT-1997.
 PR 26-MAR-1997; F00537.
 PR 26-MAR-1996; FR-003730.
 PA (VETI-) VETIGEN.

PI Drumare MF, Lenzén G, Pictet-Rouxel F, Strosberg AD:
DR MPI: 98-032136/03.
DR P-PSDB: M44932.
PT Canine beta 2 and beta 3 adrenergic receptors and coding sequences -
PT specific for identifying specific ligands and (ant)agonists to develop
PT specific treatments for obesity in dogs
PS Claim 1: Page 45-46; 79pp; French.
CC This sequence represent the coding region of the canine beta 2-adrenergic
CC receptor (RA-Ca-b2) gene. The sequence was isolated from a cDNA library
CC constructed from polyA+ RNA purified from dog brown adipose tissue cells.
CC The probe was a 600 bp fragment of the coding region of the human beta-3
CC adrenergic receptor covering the region from the initiation codon to
CC transmembrane domain 5 (TM5). The full length insert was cloned into M13
CC for sequencing using primers V30491-V30510. The sequence can then be
CC expressed e.g. in a mammalian cell, by subcloning into an expression
CC vector such as pCDNA3. The beta-2 receptor can be used in comparative
CC structure-function studies, e.g. for differential screening of ligands
CC specific for RA-Ca-b2 or RA-Ca-b3 (M44933).
SQ Sequence 2679 BP; 577 A; 736 C; 724 G; 642 T;

Query Match 83.2%; Score 15.8; DB 1; Length 2679;
Best Local Similarity 89.5%; Pred. No. 44;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 aacgcagcgcctctctgc 19
DB 184 AACCGACGCTCTTCTGC 202
|||||

RESULT 6
V80117/c
ID V80117 standard; DNA: 3334 BP.
AC V80117;
DT 15-MAR-1999 (first entry)
DE Osf2/Cbfa1 polypeptide encoding DNA.
KW Osf2/Cbfa1: osteoblast specific factor-2; CBFA1 locus; transcriptional;
KW osteogenic; gene therapy; modulator; bacterial infection; transgenic;
KW osteoblast; bone; osteocalcin; collagen; osteopontin; statoprotein; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 105..1895
FT /tag= a
FT /transl_except= (pos:1506..1508, aa:Gly)
FT /product= "Osf2/Cbfa1 polypeptide"
PN WO9554322-A1.
PD 03-DEC-1998.
PF 29-MAY-1998: U10860.
PR 24-MAR-1998: US-080189.
PR 29-MAY-1997: US-048430.
PA (TEXA) UNIV TEXAS SYSTEM.
PI Ducey P, Karsenty G;
DR MPI: 99-059837/05.
PT New nucleic acid expressing the osteoblast-specific transcription
PT factor Osf2 - useful for, e.g. treatment of osteogenic diseases, in
PT vaccines and for diagnosis.
PS Claim 5: Pages 185-187; 273pp; English.
CC This DNA encodes a native Osf2/Cbfa1 polypeptide (an osteoblast specific
CC factor-2 encoded by the CBFA1 locus). Host cells containing a vector
CC comprising a Osf2/Cbfa1 nucleic acid are used for the recombinant
CC production of the protein. The Osf2/Cbfa1 has osteoblast-specific
CC transcriptional activity (particularly for treating osteogenic diseases,
CC optionally when expressed from a gene therapy vector). Osf2/Cbfa1 is also
CC used to raise antibodies, to screen for modulators of its activity; used
CC in vaccines and to detect specific antibodies (for diagnosis of bacterial
CC infections). The Osf2/Cbfa1 polynucleotides can be used to produce
CC transgenic animals or pluripotent non-human animal cells, while their
CC fragments are used to detect Osf2/Cbfa1 genes by hybridisation, or as
CC antisense molecules or ribozymes for downregulation of gene expression.
CC Osf2/Cbfa1 polynucleotides and polypeptides are used for specific
CC transcription of osteoblast-specific genes that have an OSE2 sequence
CC element; to generate an immune response; in binding assays to detect OSE2

CC elements; for purification of such elements and to induce differentiation
CC of osteoblast progenitors for stimulating formation, growth, replacement
CC and repair of bone tissue. Antibodies, optionally, labelled, are used as
CC immunosay reagents for detecting Osf2/Cbfa1; in DNA-binding assays to
CC identify other genes to which Osf2/Cbfa1 can bind; for affinity
CC purification of Osf2/Cbfa1 and to clone related genes. Also regulatory
CC sequences (promoter and enhancer) from Osf2/Cbfa1 genes are used to
CC provide osteoblast-specific expression of homologous or heterologous
CC genes, e.g. osteocalcin, type I collagen, osteopontin and bone
CC statoprotein.
SQ Sequence 3334 BP; 840 A; 892 C; 745 G; 857 T;

Query Match 83.2%; Score 15.8; DB 1; Length 3334;
Best Local Similarity 89.5%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 aacgcagcgcctctctgc 19
DB 190 AACCGACGCTCTTCTGC 172
|||||

RESULT 7
V80182/c
ID V80182 standard; DNA: 6178 BP.
AC V80182;
DT 15-MAR-1999 (first entry)
DE Nucleic acid sequence of the Osf2 promoter and 5' UTR.
KW Osf2/Cbfa1: osteoblast specific factor-2; CBFA1 locus; transcriptional;
KW osteogenic; gene therapy; modulator; bacterial infection; transgenic;
KW osteoblast; bone; osteocalcin; collagen; osteopontin; statoprotein;
KW alternative splicing; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 6178..6178
FT WO9554322-A1.
PD 03-DEC-1998.
PF 29-MAY-1998: U10860.
PR 24-MAR-1998: US-080189.
PR 29-MAY-1997: US-048430.
PA (TEXA) UNIV TEXAS SYSTEM.
PI Ducey P, Karsenty G;
DR MPI: 99-059837/05.
PT New nucleic acid expressing the osteoblast-specific transcription
PT factor Osf2 - useful for, e.g. treatment of osteogenic diseases, in
PT vaccines and for diagnosis.
PS Claim 72: Pages 155-156; 273pp; English.
CC The invention relates to an Osf2/Cbfa1 polypeptide (an osteoblast
CC specific factor-2 encoded by the CBFA1 locus). Host cells containing a
CC vector comprising a Osf2/Cbfa1 nucleic acid are used for the recombinant
CC production of the protein. The Osf2/Cbfa1 has osteoblast-specific
CC transcriptional activity (particularly for treating osteogenic diseases,
CC optionally when expressed from a gene therapy vector). Osf2/Cbfa1 is also
CC used to raise antibodies, to screen for modulators of its activity; used
CC in vaccines and to detect specific antibodies (for diagnosis of bacterial
CC infections). The Osf2/Cbfa1 polynucleotides can be used to produce
CC transgenic animals or pluripotent non-human animal cells, while their
CC fragments are used to detect Osf2/Cbfa1 genes by hybridisation, or as
CC antisense molecules or ribozymes for downregulation of gene expression.
CC Osf2/Cbfa1 polynucleotides and polypeptides are used for specific
CC transcription of osteoblast-specific genes that have an OSE2 sequence
CC element; to generate an immune response; in binding assays to detect OSE2
CC elements; for purification of such elements and to induce differentiation
CC of osteoblast progenitors for stimulating formation, growth, replacement
CC and repair of bone tissue. Antibodies, optionally, labelled, are used as
CC immunosay reagents for detecting Osf2/Cbfa1; in DNA-binding assays to
CC identify other genes to which Osf2/Cbfa1 can bind; for affinity
CC purification of Osf2/Cbfa1 and to clone related genes. Also regulatory
CC sequences (promoter and enhancer) from Osf2/Cbfa1 genes are used to
CC provide osteoblast-specific expression of homologous or heterologous
CC genes, e.g. osteocalcin, type I collagen, osteopontin and bone
CC statoprotein. The present sequence represents the nucleic acid sequence
CC of the Osf2 promoter and 5' UTR.
SQ Sequence 6178 BP; 2195 A; 1164 C; 1148 G; 1671 T;

DB 96589 AACAGCAGCGTCTCTTGC 9660/ .

DE Flavobacterium carotenoid b

FW	Carotenoid: lycopene; beta-carotene; echinone; canthaxanthin;
KW	zeaxanthin; adonixinth; zeaxanthin; crtE;
KW	geranylgeranyl pyrophosphate synthase; GGP synthase; crtB;
KW	phytyltransferase; phytyltransferase; crtI; phytyltransferase;
KW	crtI; lycopene synthase; crtI; beta-carotene hydroxylase; ds.
OS	Flavobacterium sp. R134 WT (ATCC 21588).
PH	Key
FW	cds
FW	2. .1168
FW	/*tag- a
FW	/label- ORF-5
FW	1180. .2355
FW	/*tag- b
FW	/label- ORF-1
FW	2521. .3408
FW	/*tag- c
FW	/label- crtE
FW	/product- GGP synthase
FW	complement (3405. .4316)
FW	/*tag- d
FW	/label- crtB
FW	/product- (pre)phytyltransferase
FW	complement (4313. .5797)
FW	/*tag- e
FW	/label- crtI
FW	/product- phytyltransferase
FW	complement (5794. .6942)
FW	/*tag- f
FW	/label- crtY
FW	/product- lycopene cyclase
FW	complement (6939. .7448)
FW	/*tag- g
FW	/label- crtZ
FW	/product- beta-carotene hydroxylase
FW	complement (7767. .8315)
FW	/*tag- h
FW	/label- ORF-16
FW	8348. .8349
FW	/*tag- i
FW	/note- "bases 8348-8349 are given as nn in the
FW	specification"
FW	8539. .8540
FW	/*tag- j
FW	/note- "bases 8539-8540 are given as nn in the
FW	specification"
FW	8581
FW	/*tag- k
FW	/note- "base 8581 is given as n in the
FW	specification"
FW	8590
FW	/*tag- l
FW	/note- "base 8590 is given as n in the
FW	specification"
FW	8592
FW	/*tag- m
FW	/note- "base 8592 is given as n in the
FW	specification"
FW	8602. .8604
FW	/*tag- n
FW	/note- "bases 8602-8604 are given as nn in the
FW	specification"
FW	misc_difference
FW	misc_difference
FW	8590
FW	/*tag- o
FW	/note- "base 8590 is given as n in the
FW	specification"
FW	8592
FW	/*tag- p
FW	/note- "base 8592 is given as n in the
FW	specification"
FW	8602. .8604
FW	/*tag- q
FW	/note- "base 8602-8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- r
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- s
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- t
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- u
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- v
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- w
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- x
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- y
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- z
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- aa
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ab
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ac
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ad
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ae
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- af
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ag
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ah
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ai
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- aj
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ak
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- al
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- am
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- an
FW	/note- "base 8604 are given as nn in the
FW	specification"
FW	8604
FW	/*tag- ao
FW	/note- "base 8604 are given as nn in the

DR P-PSDB: M06518;
DR P-PSDB: M06519.
PT Flavobacterium gene sequences encoding carotenoid biosynthesis
PT enzymes - for the production of carotenoid(s), useful in foods and
PT animal feeds
PS Example 2: Fig 7: 80pp: English.
CC Genomic DNA (T45143) of Flavobacterium sp. R1534 includes genes of
CC the carotenoid biosynthesis pathway. The sequence was deduced from
CC inserts of 6 clones obtd. from genomic libraries e.g. by PCR
CC amplification (see also T45144-45) and use of partial clones to
CC screen the library. The identities of the gene products (see also
CC M06513-19 and M00871) were detd. by examining carotenoid
CC accumulation in E. coli hosts transformed with deleted variants of
CC the gene cluster. The isolated genes can be used in different
CC combinations to produce carotenoids in transformed host cells.
SO Sequence 8625 BP; 1458 A; 2898 C; 2964 G; 1295 T;

QY Query Match 81.1%; Score 15.4; DB 1; Length 8625;
Best Local Similarity 94.1%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0Y 2 acgcagcgcctctctg 18
|||||
Db 4503 ACGCAGCGCCTTCTCG 4487

RESULT 13
X13195
ID X13195 standard; DNA; 8991 BP.
AC X13195.
DT 19-MAR-1999 (first entry)
KW Enterococcus faecalis genome contig SEQ ID NO:258.
DE Enterococcus faecalis: contig; detection: Enterococcal infection;
KW vaccine: attenuation; computer readable medium; ds.
OS Enterococcus faecalis.
PN W09850555.A2.
PD 12-NOV-1998.
PF 04-MAY-1998; U08985.
PR 14-NOV-1997; US-066009.
PR 06-MAY-1997; US-044031.
PR 16-MAY-1997; US-046555.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Dillon PJ, Kunsch CA;
DR WPI: 99-045171/04.
PT New isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
PS Claim 1: Page 1239-1244; 2084pp; English.
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC X12938 to X13919 represent these nucleotide sequences which are primary
CC nucleotide sequences, also known as contigs. The computer-based system
CC can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
SO Sequence 8991 BP; 2910 A; 1478 C; 1913 G; 2683 T;

QY Query Match 81.1%; Score 15.4; DB 1; Length 8991;
Best Local Similarity 94.1%; Pred. No. 77;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

0Y 1 aacgcagcgcctctct 17
|||||

Db 8630 AATGACGCGCTTCTT 8646

RESULT 14

ID T87401 standard; DNA; 1417 BP.
 AC T87401;
 DT 07-DEC-1997 (first entry)
 DE AD4/AD3LP sequence.
 KW AD3; AD4/AD3LP; Alzheimer's disease; chromosome; missegregation;
 OS presentin-1; inhibitor; AD; trisomy 21; ss.
 FH Homo sapiens.
 FT cds
 FT 1.1132
 FT /*tag= a
 PN MO9707213-A2.
 PD 27-FEB-1997.
 PE 15-AUG-1996; U13314.
 PR 16-AUG-1995; US-002448.
 PA (HARD) HARVARD COLLEGE.
 PI Li J, Potter H.
 DR WPI: 97-165297/15.
 DR P-PSDB: W28506.
 PR Identifying genes which cause chromosome missegregation - useful for
 PT identifying causes of and treatments for diseases, e.g. Alzheimer's
 PS disease, cancer and ageing.
 PS Claim 16; Fig 1: 77pp; English.
 CC Identifying genes which cause improper chromosome segregation,
 CC screening for inhibitors of chromosome missegregation and processes
 CC caused by genes encoding chromosome missegregation promoters
 CC was exemplified using Alzheimer's disease. The sequences
 CC given in T87401 to T87426 can be used in the above methods.
 CC It is not clear from the figure legend, the figure and the
 CC disclosure of the specification which sequence of Fig 1 and Fig 28
 CC is the AD4/AD3LP or the AD3 sequence.
 SQ Sequence 1417 BP; 307 A; 385 C; 380 G; 345 T;

Query Match 78.9%; Score 15; DB 1; Length 1417;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcagcgctctcttg 18
 |||||||
 Db 1047 GGCAGCGCTTCTTG 1033

RESULT 15

ID T40031 standard; DNA; 2229 BP.
 AC T40031;
 DT 25-JUL-1997 (first entry)
 DE Human presentin-2 wild type coding sequence.
 KW Presentin-2; human; hps1-1; hps1-2; PS-2; Integral membrane protein; AD;
 KW familial Alzheimer's disease; cerebral haemorrhage; schizophrenia;
 KW depression; antibody; gene expression modulator; therapy; ss.
 OS Homo sapiens.
 FH key
 FT cds
 FT 366.1712
 FT /tag= a
 FT /product= presentin-2
 PN MO9634099-A2.
 PD 31-OCT-1996.
 PE 29-APR-1996; CA0263.
 PR 28-APR-1995; US-431048.
 PR 28-JUN-1995; US-496841.
 PR 31-JUL-1995; US-509359.
 PA (HSCR-) HSC RES & DEV LP.
 PA (UTOR) UNIV TORONTO GOVERNING COUNCIL.
 PI Fraser PE, Rommens JM, St George-Hyslop PH;
 DR WPI: 96-497631/49.
 DR P-PSDB: W05762.
 PT New presentin genes - useful for diagnosis, therapy and drug

PR screening of familial Alzheimer's disease, cerebral disorders, etc.
 PS Claim 8; Page 148-150; 178pp; English.
 CC This sequence represents the coding sequence for the human presentin-2.
 CC T40028 and T40029 represent the coding sequences for the two different
 CC forms of wild type human presentin-1 (PS-1). The form represented by
 CC T40029 results from alternate splicing of the genomic DNA sequence.
 CC presentin-1 are a family of highly conserved integral membrane proteins
 CC with a common structural motif, common alternate splicing patterns, and
 CC common mutational hot spot regions. Mutations in PS genes are implicated
 CC in familial Alzheimer's disease (AD) and possibly other diseases such as
 CC cerebral haemorrhage, schizophrenia, depression etc., so detection of
 CC mutations in these sequences can be used for diagnosis of these diseases.
 CC The encoded proteins, or vectors that express them or containing
 CC antisense sequences, antibodies selective for mutant forms of the encoded
 CC proteins (such as W05736) and modulators of PS gene expression are useful
 CC potentially useful for treatment of AD etc. Transgenic animals are useful
 CC as models for drug screening. The antibodies can also be used e.g. for
 CC affinity purification and in immunosays.
 SQ Sequence 2229 BP; 481 A; 579 C; 633 G; 521 T;

Query Match 78.9%; Score 15; DB 1; Length 2229;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcagcgctctcttg 18
 |||||||
 Db 1606 GGCAGCGCTTCTTG 1592

Search completed: September 12, 2000, 23:07:17
 Job time: 4099 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:04:28 : Search time 112.94 Seconds
(without alignments)
23.141 Million cell updates/sec

Title: US-09-542-718-2

Perfect score: 19
Sequence: 1 aacggcagcgctctctgc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/5C.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/5D.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/6.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/PCITUS.COMB.seq: *
7: /cgn2_6/ptodata/1/lna/Backfile1.seq: *

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	19	100.0	1242	6	PCT-US91-00909-3
2	15.4	81.1	1400	3	US-08-001-078A-2
3	15.4	81.1	1400	4	US-08-463-218-2
4	15.4	81.1	1400	6	PCT-US94-00253-2
5	15	78.9	1417	4	US-08-875-972-1
6	15	78.9	2229	5	US-08-888-077A-18
7	15	78.9	2276	4	US-08-875-972-28
8	15	78.9	2285	3	US-08-967-101-136
9	15	78.9	2285	4	US-08-592-541-136
10	14.4	75.8	1113	5	US-08-651-136C-65
11	14.4	75.8	1912	1	US-08-270-013B-1
12	14.4	75.8	1912	2	US-08-838-418-1
13	14.4	75.8	3845	4	US-08-220-240A-4
14	14.4	74.7	501	5	US-08-384-106A-8
15	14.2	74.7	501	6	PCT-US96-01643-8
16	14.2	74.7	555	5	US-08-384-106A-7
17	14.2	74.7	555	5	US-08-384-106A-19
18	14.2	74.7	555	6	PCT-US96-01643-7
19	14.2	74.7	706	6	PCT-US96-05252-1
20	14.2	74.7	1254	6	PCT-US91-00909-1
21	14.2	74.7	2455	1	US-08-073-807A-1
22	14.2	74.7	3425	3	US-08-768-147B-1
23	14.2	74.7	4465	1	US-08-180-195-1
24	14.2	74.7	4465	1	US-08-477-329-1
25	14.2	74.7	4465	3	US-08-475-458-1
26	14.2	74.7	4465	5	US-08-980-400-1

27	14.2	74.7	5427	1	US-08-168-917-1	Sequence 1, Appl
28	14.2	74.7	5427	3	US-08-460-510-1	Sequence 1, Appl
29	14.2	74.7	5427	5	US-08-460-490-1	Sequence 1, Appl
30	14.2	74.7	5427	3	US-08-462-728-3	Sequence 3, Appl
31	14.2	74.7	5427	6	PCT-US92-00730-1	Sequence 1, Appl
32	14.2	74.7	5427	6	PCT-US92-00662-1	Sequence 1, Appl
33	13.8	72.6	835	4	US-08-557-309B-19	Sequence 19, Appl
34	13.8	72.6	835	5	US-08-834-306-19	Sequence 19, Appl
35	13.8	72.6	853	3	US-08-695-736-2	Sequence 2, Appl
36	13.8	72.6	1202	4	US-08-428-414A-1	Sequence 1, Appl
37	13.8	72.6	1245	1	US-07-887-072B-1	Sequence 1, Appl
38	13.8	72.6	1245	2	US-08-466-444-1	Sequence 1, Appl
39	13.8	72.6	1401	4	US-08-744-779A-1	Sequence 1, Appl
40	13.8	72.6	1401	4	US-08-862-531-1	Sequence 1, Appl
41	13.8	72.6	1894	6	PCT-US93-06404-3	Sequence 1, Appl
42	13.8	72.6	2692	1	US-07-932-454A-2	Sequence 3, Appl
43	13.8	72.6	3182	1	US-08-188-582-12	Sequence 12, Appl
44	13.8	72.6	3182	1	US-08-646-715-12	Sequence 12, Appl
45	13.8	72.6	3453	6	PCT-US92-05401-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
Sequence 3, Application PC/TUS9100909
GENERAL INFORMATION:
APPLICANT: Siedziewski, Andrzej Z.
TITLE OF INVENTION: Methods of Producing Hybrid G
TITLE OF INVENTION: Proteins-Coupled Receptors
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: United States
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/00909
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.408PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-602-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1242 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1242
PCT-US91-00909-3
Query Match 100.0% Score 19; DB 6; Length 1242;
Best Local Similarity 100.0% Pred. No. 0.83;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aacgcagcgcctctctg 19
|||
Db 16 AACGCAGCGCCTCTCTGC 34

RESULT 2

US-08-001-078A-2/c
Sequence 2, Application US/08001078A
Patent No. 5872094
GENERAL INFORMATION:
APPLICANT: Goettrich, Paul F.
APPLICANT: Tondravil, M., Mehرداد
APPLICANT: Binette, Francois
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/001,078A
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-001-078A-2

Query Match 81.1%; Score 15.4; DB 3; Length 1400;

Best Local Similarity 94.1%; Pred. No. 42;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 acgcagcgcctctctg 18
|||
Db 1212 ACTGCAGCGCCTCTCTG 1196

US-08-463-218-2/c

Sequence 2, Application US/08463218

Patent No. 5986052

GENERAL INFORMATION:

APPLICANT: Goettrich, Paul F.

APPLICANT: Tondravil, Mehرداد

APPLICANT: Binette, Francois

TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 STATE STREET, Suite 510

CITY: BOSTON

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,218

FILING DATE: 05-JUNE-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/001,078

FILING DATE: 06-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Paul L.

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: MGP-008DV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1400 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear
US-08-463-218-2

Query Match 81.1%; Score 15.4; DB 4; Length 1400;

Best Local Similarity 94.1%; Pred. No. 42;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 acgcagcgcctctctg 18
|||
Db 1212 ACTGCAGCGCCTCTCTG 1196

US-08-463-218-2

RESULT 4

PCT-US94-00253-2/c
Sequence 2, Application PC/TUS9400253
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
NUMBER OF SEQUENCES: 2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00253
PRIOR APPLICATION DATA: US 08/001,078
APPLICATION NUMBER: US 08/001,078
FILING DATE: 06-JAN-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US94-00253-2

Query Match 81.1%; Score 15.4; DB 6; Length 1400;

Best Local Similarity 94.1%; Pred. No. 42;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 acgcagcgcctctctg 18
|||
Db 1212 ACTGCAGCGCCTCTCTG 1196

US-08-463-218-2

```
RESULT 5
US-08-875-972-1/c
; Sequence 1, Application US/08875972
; Patent No. 5985564
; GENERAL INFORMATION:
; APPLICANT: Huntington Potter and Jinhue Li
; TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
; NUMBER OF SEQUENCES: 29
; CHROMOSOME NON-DISJUNCTION
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,972
; FILING DATE: 08-AUG-97
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,448
; FILING DATE: 16-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan Esq., Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: H095-03PA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1129
; US-08-875-972-1

Query Match 78.9%; Score 15; DB 4; Length 1417;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcagcgcctcttg 18
|||||
Db 1047 ggcagcgcctcttg 1033

RESULT 6
US-08-888-077A-18/c
; Sequence 18, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USDS THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESSES:
; ADDRESS: LERNER, DAVID, LITTENBERG, KRUMHOLTZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
```

```
CITY: WESTFIELD
STATE: NJ
COUNTRY: USA
ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2229 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 366..1712
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2226
; OTHER INFORMATION: /note="hps2"
; US-08-888-077A-18

Query Match 78.9%; Score 15; DB 5; Length 2229;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ggcagcgcctcttg 18
|||||
Db 1606 ggcagcgcctcttg 1592

RESULT 7
US-08-875-972-28/c
; Sequence 28, Application US/08875972
; Patent No. 5985564
; GENERAL INFORMATION:
; APPLICANT: Huntington Potter and Jinhue Li
; TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
; NUMBER OF SEQUENCES: 29
; CHROMOSOME NON-DISJUNCTION
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millita Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/875,972
; FILING DATE: 08-AUG-97
; CLASSIFICATION: 435
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,448
FILING DATE: 16-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Granahan Esq., Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: H095-03PA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2276 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-875-972-28

Query Match 78.9%; Score 15; DB 4; Length 2276;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcagcgccctctg 18
|||||
Db 1648 GGCAGCGCCTTCTTG 1634

RESULT 8
US-08-967-101-136/c
Sequence 136, Application US/08967101
Patent No. 5840540
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,101
FILING DATE: 10-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/592,541
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 2285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-967-101-136

Query Match 78.9%; Score 15; DB 3; Length 2285;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcagcgccctctg 18
|||||
Db 1606 GGCAGCGCCTTCTTG 1592

RESULT 9
US-08-592-541-136/c
Sequence 136, Application US/08592541
Patent No. 5986054
GENERAL INFORMATION:
APPLICANT: ST. GEORGE-HYSLOP, PETER H
APPLICANT: ROMMENS, JOHANNA M
APPLICANT: FRASER, PAUL E
TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
TO ALZHEIMER'S DISEASE
NUMBER OF SEQUENCES: 183
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: High Street Tower - 125 High Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,541
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7100
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 136:
SEQUENCE CHARACTERISTICS:
LENGTH: 2285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-592-541-136

Query Match 78.9%; Score 15; DB 4; Length 2285;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggcagcgccctctg 18
|||||
Db 1606 GGCAGCGCCTTCTTG 1592

RESULT 10
US-08-651-136C-63
Sequence 63, Application US/08651136C
Patent No. 6001639
GENERAL INFORMATION:
APPLICANT: Schulein, Martin
APPLICANT: Andersen, Lene N.
APPLICANT: Lassen, Soren F.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Lange, Lene
APPLICANT: Nielsen, Rudy I.
APPLICANT: Ihara, Michiko

TELEFAX: (312) 616-5700
TELEX: (25)3533
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1912 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus stearotheophilus
STRAIN: ATCC12016
US-08-838-418-1

Query Match 75.8%; Score 14.4; DB 2; Length 1912;
Best Local Similarity 93.8%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 aacggcagccttct 16
|||||

Db 1361 AACGGCAGCCTTCT 1376

RESULT 13
US-08-220-240A-4
Sequence 4, Application US/08220240A
Patent No. 5955291
GENERAL INFORMATION:
APPLICANT: Altalo, Karl
APPLICANT: Matikainen, Marja-Terttu
APPLICANT: Partanen, Juha
APPLICANT: Mäkelä, Tomi
APPLICANT: Korhonen, Jaana
TITLE OF INVENTION: ANTIBODIES RECOGNIZING THE RECEPTOR
TITLE OF INVENTION: TYROSINE KINASE AND USES THEREOF
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220,240A
FILING DATE: 29-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FI93/00006
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/817,800
FILING DATE: 09-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/167,453
FILING DATE: 15-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gass, David A.
REGISTRATION NUMBER: 38,153
REFERENCE/DOCKET NUMBER: 29151/31958
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 3845 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 37..3450
US-08-220-240A-4

Query Match 75.8%; Score 14.4; DB 4; Length 3845;
Best Local Similarity 93.8%; Pred. No. 1.3e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 ggcagcgccttcttgc 19
|||||

Db 646 GGCAGCGCCTTCTTGC 661

RESULT 14
US-08-384-106A-8/C
Sequence 8, Application US/08384106A
Patent No. 6033847
GENERAL INFORMATION:
APPLICANT: Sheer Ph.D., Charles J.
APPLICANT: Downing M.D., James
APPLICANT: Hirai Ph.D., Hiroshi
APPLICANT: Okuda, Tsukasa
TITLE OF INVENTION: Inkc-p18 and Inkc4-p19, Inhibitors of
TITLE OF INVENTION: Cyclin Dependent Kinases CDK4 and CDK6, and Uses Thereof
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX
STREET: 1100 New York Ave., N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,106A
FILING DATE: 06-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656.0500000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..501
US-08-384-106A-8

Query Match 74.7%; Score 14.2; DB 5; Length 501;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 aacggcagccttcttgc 19

DB 143 ACCTGACGCCGCTTGC 125

RESULT 15

PCT-US96-01643-8/C
Sequence 8, Application PC/TUS9601643

GENERAL INFORMATION:

APPLICANT: St. Jude Children's Research Hospital
TITLE OF INVENTION: Ink4c-p18 and Ink4d-p19, Inhibitors of
CYCLIN DEPENDENT KINASES CDK4 and CDK6, and uses thereof
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01643
FILING DATE: 06-FEB-1996

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/384,106
FILING DATE: 06-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Fox, Samuel L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0656,050PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 501 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: p19
FEATURE:
NAME/KEY: CDS
LOCATION: 1..501

PCT-US96-01643-8

Query Match 74.7%; Score 14.2; DB 6; Length 501;
Best Local Similarity 84.2%; Pred. No. 1.4e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 aaaggagcgccttcttc 19
DB 143 ACCTGACGCCGCTTGC 125

Search completed: September 12, 2000, 23:04:32
Job time: 3939 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 22:46:38 ; Search time 1893.64 Seconds
(without alignments)
44.246 Million cell updates/sec

Title: US-09-542-718-2

Perfect score: 19

Sequence: 1 aacggcagcgccctctgc 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*
110: gb_est91:*
111: gb_est92:*
112: gb_est93:*
113: gb_est94:*
114: gb_est95:*
115: gb_est96:*
116: gb_est97:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	100.0	377	80	C18250 C18250 Homo sapiens
2	17.4	91.6	195	40	A1475044 tm62603.x
3	17.4	91.6	283	64	AM139725 UI-H-B11
4	17.4	91.6	369	35	AM128872 qf1608.s
5	17.4	91.6	442	36	AM199676 q160209.x
6	17.4	91.6	460	44	AM1828151 WK3210.x
7	17.4	91.6	527	46	AM1939616 UI-M-A10
8	16	84.2	422	23	AA234907 z338f03.x
9	16	84.2	423	22	AA150236 z103c01.r
C 10	15.8	83.2	195	86	N27820 YX5410.r
C 11	15.8	83.2	195	86	R86608 RABEST229T
C 12	15.8	83.2	197	39	AM143640 MD29112.x
C 13	15.8	83.2	234	89	R86588 RABEST163T
C 14	15.8	83.2	235	89	R86566 RABEST140T
C 15	15.8	83.2	275	89	R86542 RABEST103T
C 16	15.8	83.2	279	89	R86528 RABEST084T
C 17	15.8	83.2	323	73	AM491133 UI-M-BH3
C 18	15.8	83.2	331	89	R86510 RABEST202T
C 19	15.8	83.2	345	48	AM046197 AU046197
C 20	15.8	83.2	349	63	AM036224 EST278222
C 21	15.8	83.2	364	45	AM1835798 UI-M-A10
C 22	15.8	83.2	386	45	AM1846942 UI-M-AK1
C 23	15.8	83.2	393	45	AM1843043 UI-M-AK1
C 24	15.8	83.2	397	45	AM1842269 UI-M-A11
C 25	15.8	83.2	397	45	AM1848899 UI-M-AJ1
C 26	15.8	83.2	398	45	AM1836035 UI-M-AJ1
C 27	15.8	83.2	398	90	W15007 mb29h12.r1
C 28	15.8	83.2	428	63	AM048256 UI-M-BH1
C 29	15.8	83.2	430	73	AM493285 UI-M-BH3
C 30	15.8	83.2	435	45	AM1844706 UI-M-A11
C 31	15.8	83.2	442	39	AM1390851 mb29h12.Y
C 32	15.8	83.2	456	63	AM046979 UI-M-BH1
C 33	15.8	83.2	472	45	AM1836840 UI-M-AJ0
C 34	15.8	83.2	472	91	W31412 z291c01.r1
C 35	15.8	83.2	473	45	AM1836000 UI-M-AJ0
C 36	15.8	83.2	507	114	AM0803246 HS_3154_A
C 37	15.8	83.2	573	73	AM533116 UI-R-BU0
C 38	15.8	83.2	593	69	AM186808 BNLGH1524
C 39	15.8	83.2	650	38	AM1353264 zeh0256.s
C 40	15.8	83.2	694	48	AM070300 AU070300
C 41	15.8	83.2	713	48	AU070299 AU070299
C 42	15.8	83.2	744	33	AA980445 ua4102.r
C 43	15.4	81.1	98	21	AA114801 mml6c09.r
C 44	15.4	81.1	123	79	AM677465 DGI_7_F12
C 45	15.4	81.1	334	45	AM1876306 WK73f09.Y

ALIGNMENTS

RESULT 1
LOCUS C18250/c
DEFINITION C18250 Human placenta cDNA (TFujiwara) Homo sapiens cDNA clone
GEN:559E06 5', mRNA sequence.

ACCESSION C18250
VERSION C18250.1 GI:1579852
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Fujiwara, T., Hirano, H., Hishigaki, H., Horie, M., Kawai, A., Kuga, Y., Kyushiki, H., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Suzuki, M., Takachi, A., Takeda, S., Watanabe, T., Maekawa, H., Nakamura, Y. and Takahashi, E.
1 (bases 1 to 377)

TITLE
JOURNAL
COMMENT Contact: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawauchi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.

FEATURES
source Location/Qualifiers

1..377
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GEN-559E06"
/clone_11b="Human placenta cDNA (TFujiwara)"
/tissue_type="placenta"
BASE COUNT 91 a 113 c 101 g 72 t
ORIGIN

Query Match 100.0%; Score 19; DB 80; Length 377;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aacggcagcgctctctgc 19
|||||
Db 342 AACGCGACGCGCTCTGCG 324

RESULT 2
LOCUS A1475044
DEFINITION A1475044 195 bp mRNA EST 14-APR-1999
similar to contains element TARI repetitive element ;, mRNA
sequence.

ACCESSION A1475044
VERSION A1475044.1 GI:4328089
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 195)
NCBI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAR/BRGAR), Tumor Gene Index

JOURNAL
COMMENT Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert length: 1052 Std Error: 0.00
Seq primer: -40UP from Gibco

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2718667"
/clone_lib="NCI_CGAP_Sub3"
/_lib_host="DH10B (Life Technologies)"
/notice_vector="p773D-Pac (Pharmacia)" with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The
NCI_CGAP_Sub3 library is a subtracted library derived from
the NCI_CGAP_Sub1 library, which is a subtracted library
derived from B1. B1 constitutes a mixture of 21
normalized or subtracted NCI_CGAP libraries
NCI_CGAP_C04, NCI_CGAP_P122, NCI_CGAP_P128, NCI_CGAP_C010,

```

BASE COUNT
CRIGIN

34 a 128 c 82 g 39 t

NCI_CGAP_C016, NCI_CGAP_K1d5, NCI_CGAP_K1d12,
NCI_CGAP_K1d3, NCI_CGAP_K1d11, NCI_CGAP_LYM2,
NCI_CGAP_Br2, NCI_CGAP_C08, NCI_CGAP_C111, NCI_CGAP_L612,
NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:

NCI_CGAP_K1d3 pool 1 LLM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clonoids 132376-133391),
1456008-1456775, 1500552-1502855); NCI_CGAP_K1d5 pool 1
LLM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids
1333912-1325831, 1471368-1472903, 1492104-1493355);
NCI_CGAP_Lu5 pool 1 LLM 3575-3582, 3851-3854 (IMAGE
Clonoids 1414920-1417991, 1520904-1524239); NCI_CGAP_GC4
pool 1 LLM 3164-3167, 3716-3720, 3733-3735 (IMAGE
Clonoids 1257096-1258631, 1469064-1470984),
1475592-1476743); NCI_CGAP_Pr2 pool 1 LLM 2457-2459,
2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759,
1101192-1101959, 1217928-1220615); NCI_CGAP_C01 pool 1
LLM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255,
1144584-1145351). Subtraction was performed as previously
described [Bonaldo, Lennon & Soares (1996)]. Normalization
and Subtraction: Two Approaches To Facilitate Gene
Discovery. Genome Research 6, 791-806.;

TAG_L1B-NCI_CGAP_Pr22; TAG_TISSUE-prostate; TAG_SEO-NAOCT*

Query Match	91.6%	Score	17.4	DB	64	Length	283
Best Local Similarity	94.7%	Pred. No.	1.1e+02				
Matches	18	Conservative	0	Mismatches	1	Indels	0
						Gaps	0
QY	1	aacggcagcgccctcttcgc	19				
Db	196	AACGGCACCGCCCTTCCTGC	214				
RESULT	4						
LOCUS	A1128872						
DEFINITION	A1128872	369 bp	mRNA	EST	13-FEB-1999		
	glt6g98.g.1	NCI_CGAP_Brn25	Homo sapiens	CDNA clone IMAGE:1750238	3',		
ACCESSION	A1128872						
VERSION	A1128872.1	GI:3597386					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 369)						
AUTHORS	NCI/NCDS-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.					
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project						
	(CGAP/BRCA), Tumor Gene Index						
	Unpublished (1998)						
JOURNAL	On May 18, 1998 this sequence version replaced gi:3136833.						
COMMENT	Contact: Robert Strausberg, Ph.D.						
	Tel: (301) 496-1550						
	Email: Robert.Strausberg@nih.gov						
	Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.						
	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.						
	CDNA Library Arrayed by: Greg Lannon, Ph.D.						
	DNA sequencing by: Washington University Genome Sequencing Center						
	clone distribution: NCI-CGAP clone distribution information can be						
	found through the I.M.A.G.E. Consortium/LNLB at:						
	www-bio.llnl.gov/bdrrp/image/image.html						
	Insert Length: 1035	Std Error: 0.00					
	Seq primer: -40ml3 fwd. ET from Amersham						

BASE COUNT 57 a 193 c 151 g 59 t
ORIGIN

Query Match 91.6%; Score 17.4; DB 44; Length 460;
Best Local Similarity 94.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 aacggcagcgctctctgc 19
|||||
Db 154 AACGCACGCCCTTCTCTGC 172

RESULT 7
AI939616 527 bp mRNA EST 13-DEC-1999
LOCUS tm62e03.x5 NCI-CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2162716 3';
DEFINITION mRNA sequence.
ACCESSION AI939616
VERSION AI939616.1 GI:5678486
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 527)
NCI/NIHNS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index
Unpublished (1998)
On Oct 13, 1998 this sequence version replaced gi:3738881.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

JOURNAL
COMMENT

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: National Cancer Institute / National
Institute of Neurological Disorders and Stroke, Brain Tumor Genome
Anatomy Project (CGAP/BRGAP), Tumor Gene Index
This read has been verified (found to hit its original self in the
correct orientation) Std Error: 0.00
Insert Length: 1052
Seq primer: -400P from GIBCO
High quality sequence stop: 459.
Location/Qualifiers
1. 527

FEATURES
source
1. 527
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2162716"
/clone_lib="NCI-CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: Brain; Vector: p7773D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACCATCTGAGGTGGAGCGCCGATAGGCTTTTCTTTTCTTTTCTTTT
T 3'); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7773 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldi."

BASE COUNT 65 a 212 c 180 g 70 t
ORIGIN

Query Match 91.6%; Score 17.4; DB 46; Length 527;
Best Local Similarity 94.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 aacggcagcgctctctgc 19
|||||
Db 171 AACGCACGCCCTTCTCTGC 189

RESULT 8
AA234907 422 bp mRNA EST 07-AUG-1997
LOCUS zs38f03.r1 Soares.NHMPu-S1 Homo sapiens cDNA clone IMAGE:687485 5';
DEFINITION similar to TR:G1213518 G1213518 ALG3 ;, mRNA sequence.
ACCESSION AA234907
VERSION AA234907.1 GI:1859605
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 422)
Hillier, L., Lennon, G., Becker, M., Bonaldi, M.F., Chapell, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Pavello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacey, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, D., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 825 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 387.
Location/Qualifiers
1. 422

JOURNAL
COMMENT

This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: National Cancer Institute / National
Institute of Neurological Disorders and Stroke, Brain Tumor Genome
Anatomy Project (CGAP/BRGAP), Tumor Gene Index
This read has been verified (found to hit its original self in the
correct orientation) Std Error: 0.00
Insert Length: 825
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 387.
Location/Qualifiers
1. 422

FEATURES
source
1. 422
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="GDB:5591635"
/clone="IMAGE:687485"
/clone_lib="Soares.NHMPu-S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p7773D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NHM, pregnant uterus
NHMPU, and fetal heart NBH19W) were mixed, and as circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 88 a 113 c 107 g 114 t
ORIGIN

Query Match 84.2%; Score 16; DB 23; Length 422;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 cgcgcagcgcctcttg 18
 |||||||
 Db 140 CGGCAGCGCCTTCTTG 125

RESULT 9
 AA150236/c 423 bp mRNA EST 14-MAY-1997
 LOCUS N27820/c
 DEFINITION yx54h10.r1 Soares-pregnant_uterus_NbHPU Homo sapiens cDNA clone
 IMAGE:491232 5' similar to TR:G1213518 G1213518 ALG3 ; mRNA
 sequence.

ACCESSION AA150236
 VERSION AA150236.1 GI:1721824
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 423)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 AUTHORS Hillier, L., Lennon, G., Becker, M., Donald, M.F., Chiappelli, B.,
 Chissee, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
 Madis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevasakis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
 and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478

COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 825 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 401.

FEATURES

Location/Qualifiers

1..423
 /organism="Homo sapiens"
 /db_xref="GDB:3805819"
 /db_xref="taxon:9606"
 /clone="IMAGE:491232"
 /clone_1ib="Soares_pregnant_uterus_NbHPU"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pT73-Pac; Site:1; Not I -
 Site:2; Eco RI; 1st strand cDNA was primed with a Not I -
 oligo (dT) primer [5',
 AACTGAGAGATTCGCGCGCGCCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by M. Fatima Bonaldo."

BASE COUNT 82 a 112 c 104 g 118 t
 ORIGIN

Query Match 84.2%; Score 16; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 cgcgcagcgcctcttg 18
 |||||||
 Db 127 CGGCAGCGCCTTCTTG 112

RESULT 10
 N27820 498 bp mRNA EST 30-DEC-1995
 LOCUS yx54h10.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
 DEFINITION IMAGE:265603 5', mRNA sequence.

ACCESSION N27820
 VERSION N27820.1 GI:1142301
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 498)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
 Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and
 Wilson, R.

TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT On Oct 18, 1995 this sequence version replaced gi:1023122.

CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu
 High quality sequence stops: 220
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: T7
 High quality sequence stop: 220.

FEATURES

Location/Qualifiers

1..498
 /organism="Homo sapiens"
 /db_xref="GDB:3875245"
 /db_xref="taxon:9606"
 /clone="IMAGE:265603"
 /clone_1ib="Soares melanocyte 2NbHM"
 /sex="Male"
 /tissue_type="melanocyte"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pT73D (Pharmacia) with a modified
 polylinker; Site:1; Not I; Site:2; Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAGTGCAGCGCGCCTTTTCTTTTCTTTT 3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library constructed by Bento Soares and
 M. Fatima Bonaldo. RNA from normal foreskin melanocytes
 (FS374) was kindly provided by Dr. Anthony P. Albin."

BASE COUNT 99 a 125 c 137 g 131 t
 ORIGIN

Query Match 84.2%; Score 16; DB 86; Length 498;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 cgcgcagcgcctcttg 18
 |||||||
 Db 210 CGGCAGCGCCTTCTTG 195

RESULT 11
 R86608 195 bp mRNA EST 17-AUG-1995
 LOCUS RABEST229T Rabbit osteoclast, Dennis Sakai Oryctolagus cuniculus
 DEFINITION CDNA clone pRABOC229 5' similar to cofilin, mRNA sequence.
 ACCESSION R86608
 VERSION R86608.1 GI:947286

KEYWORDS
EST.
SOURCE Oryctolagus cuniculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE
1 (bases 1 to 195)
Sakai, D., Tong, H.-S. and Minkin, C.
Osteoclast Molecular Phenotyping by Random cDNA Sequencing
Bone 17 (2), 111-119 (1995)

AUTHORS
JOURNAL
MEDLINE
COMMENT
On May 10, 1995 this sequence version replaced g1:805895.
Other ESTs: RABEST038T, RABEST045T, RABEST084T, RABEST103T,
RABEST140T, RABEST163T, RABEST202T
Contact: Sakai D
Basic Sciences
University of Southern California
USC School of Dentistry, 925 West 34th Street, DEN-4220, Los
Angeles, CA 90089-0641
Tel: 2137405563
Fax: 2137407560
Email: sakai@molbio.usc.edu
Insert Length: 606 Std Error: 0.00
Seq primer: 17 promoter
High quality sequence stop: 279.

FEATURES
source
1..195
Location/Qualifiers
/organism="Oryctolagus cuniculus"
/strain="New Zealand White"
/db_xref="taxon:9986"
/clone="PRABOC229"
/clone_1lb="Rabbit Osteoclast, Dennis Sakai"
/lab_host="E. coli DH12S"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI;
Poly(A)+ RNA was purified from a 97% pure population of
osteoclasts prepared from the long bones of 10 day old
rabbits. First strand cDNA was synthesized by priming
with an oligo(dT)-NotI anchor-primer and second strand
cDNA was synthesized by replacement synthesis as described
by Gubler and Hoffman (Gene 25:283, 1983). Following the
addition of SalI adapters and NotI digestion, the cDNA was
cloned between the SalI (5c) and NotI (3c) sites of the
pSPORT1 (BRL) plasmid vector."

BASE COUNT
46 a 44 c 46 g 48 t 11 others

ORIGIN

Query Match
Best Local Similarity 83.2%; Score 15.8; DB 89; Length 195;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY
1 aacgcagcgcctcttcg 19
|||||
Db 39 AACGACGCCCTTCTGC 21

RESULT 12
AI413640 197 bp mRNA EST 09-FEB-1999
LOCUS mb29h12.x1 Soares mouse p3MNF19.5 Mus musculus cDNA clone
DEFINITION IMAGE:330887 3', mRNA sequence.
ACCESSION AI413640
VERSION AI413640.1 GI:4257144
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 197)
Matta, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.

TITLE
JOURNAL
COMMENT
The Mashu-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Maria M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
High quality sequence stop: 187.

FEATURES
source
1..197
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:330887"
/clone_1lb="Soares mouse p3MNF19.5"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: p7T3D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTATCAATCTGAGAGGAGCGCGCATTTTTTTTTTTT 3']
TGTATCAATCTGAGAGGAGCGCGCATTTTTTTTTTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."

BASE COUNT
26 a 77 c 58 g 36 t

ORIGIN

Query Match
Best Local Similarity 83.2%; Score 15.8; DB 39; Length 197;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY
1 aacgcagcgcctcttcg 19
|||||
Db 121 AACGACGCCCTTCTGC 139

RESULT 13
R86588 234 bp mRNA EST 17-AUG-1995
LOCUS RABEST163T Rabbit Osteoclast, Dennis Sakai Oryctolagus cuniculus
DEFINITION CDNA clone PRABOC163 5' similar to cofilin, mRNA sequence.
ACCESSION R86588
VERSION R86588.1 GI:947242
KEYWORDS EST.
SOURCE Oryctolagus cuniculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
1 (bases 1 to 234)
Sakai, D., Tong, H.-S. and Minkin, C.
Osteoclast Molecular Phenotyping by Random cDNA Sequencing
Bone 17 (2), 111-119 (1995)

REFERENCE
1 (bases 1 to 234)
Sakai, D., Tong, H.-S. and Minkin, C.
Osteoclast Molecular Phenotyping by Random cDNA Sequencing
Bone 17 (2), 111-119 (1995)

AUTHORS
JOURNAL
MEDLINE
COMMENT
On May 10, 1995 this sequence version replaced g1:805875.
Other ESTs: RABEST038T, RABEST045T, RABEST084T, RABEST103T,
RABEST140T, RABEST202T, RABEST229T
Contact: Sakai D
Basic Sciences
University of Southern California
USC School of Dentistry, 925 West 34th Street, DEN-4220, Los
Angeles, CA 90089-0641
Tel: 2137405563
Fax: 2137407560
Email: sakai@molbio.usc.edu

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: September 12, 2000, 23:02:47 ; Search time 930.45 Seconds
(without alignments)
36.444 Million cell updates/sec

Title: US-09-542-718-3

Perfect score: 19

Sequence: 1 acatgacgacgccatgcc 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pri:*
10: gb_pri2:*
11: gb_pri3:*
12: gb_ro:*
13: gb_sts:*
14: gb_ay:*
15: gb_un:*
16: em_fun:*
17: em_hum1:*
18: em_hum2:*
19: em_in:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_ay:*
29: em_un:*
30: em_vl:*
31: gb_hc91:*
32: gb_hc92:*
33: gb_in1:*
34: gb_in2:*
35: em_ba1:*
36: em_ba2:*
37: em_hum3:*
38: em_hum4:*
39: gb_pr4:*
40: gb_hc93:*
41: gb_hc94:*
42: gb_hc95:*
43: gb_hc96:*

44: gb_hc97:*
45: em_hc91:*
46: em_hc92:*
47: em_hc93:*
48: em_hum5:*
49: gb_pl3:*
50: gb_pr5:*
51: gb_hc98:*
52: gb_hc99:*
53: gb_hc10:*
54: gb_hc11:*
55: gb_hc12:*
56: gb_hc13:*
57: gb_hc14:*
58: gb_in3:*
59: gb_hc15:*
60: gb_hc16:*
61: gb_hc17:*
62: em_hc94:*
63: em_hc95:*
64: em_hc96:*
65: em_hc97:*
66: em_hum6:*
67: gb_hc18:*
68: gb_hc19:*
69: gb_hc20:*
70: gb_hc21:*
71: gb_hc22:*
72: gb_hc23:*
73: gb_hc24:*
74: gb_hc25:*
75: gb_hc26:*
76: gb_hc27:*
77: gb_hc28:*
78: gb_hc29:*
79: gb_hc30:*
80: gb_hc31:*
81: gb_vl1:*
82: gb_vl2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	19	100.0	1242	11 AF022953	AF022953 Homo sapi
C 2	19	100.0	1242	11 AF022954	AF022954 Homo sapi
C 3	19	100.0	1242	11 AF022955	AF022955 Homo sapi
C 4	19	100.0	1242	11 AF022956	AF022956 Homo sapi
C 5	19	100.0	1286	39 AF020305	AF020305 Homo sapi
C 6	19	100.0	1290	39 AF169225	AF169225 Homo sapi
C 7	19	100.0	1290	39 AF203386	AF203386 Homo sapi
C 8	19	100.0	1298	3 CFU73206	U73206 Canis famli
C 9	19	100.0	1320	11 MACB2AR	L38905 Macaca mula
C 10	19	100.0	1536	3 AF192345	AF192345 Felis cat
C 11	19	100.0	1948	3 CFB2AR	X94608 C. familliar
C 12	19	100.0	1970	9 HSBAR	X04827 Human mRNA
C 13	19	100.0	2305	9 HSBAR	Y00106 Human gene
C 14	19	100.0	2679	5 A65720	A65720 Sequence 1
C 15	19	100.0	3451	10 HUMADRBR	M15169 Human beta-
C 16	19	100.0	3458	10 HUMADRBR	J02960 Human beta-
C 17	19	100.0	132858	43 AC011354	AC011354 Homo sapi
C 18	19	100.0	134419	43 AC011334	AC011334 Homo sapi
C 19	17.4	91.6	2032	3 BFB2ADREC	286037 B. taurus mr
C 20	17.4	91.6	5288	3 PTGB2AR	AF000134 Sus scrof
C 21	17.4	91.6	196541	61 AC019007	AC019007 Homo sapi
C 22	17.4	91.6	233018	78 AC019279	AC019279 Homo sapi
C 23	16.4	86.3	1694	8 ZMA011615	AJ011615 Zea mays
C 24	16.4	86.3	2676	1 ECM5YB	X59939 E. coli ORF1

C 25	16.4	86.3	3873	12	MMUJ3305	AJ223305 Mus muscu
C 26	16.4	86.3	6064	12	MMU223834	AJ223834 Mus muscu
C 27	16.4	86.3	10980	2	AE000206	AE000206 Escherich
C 28	16.4	86.3	19201	1	D90742	D90742 Escherichia
C 29	16.4	86.3	19418	1	D90743	D90743 Escherichia
C 30	16.4	86.3	25656	43	AC020181	AC020181 Drosophill
C 31	16.4	86.3	98491	12	AF132039	AF132039 Mus muscu
C 32	16.4	86.3	128447	41	AC006490	AC006490 Drosophill
C 33	16.4	86.3	134199	1	SYCSLR	D64004 Synecocyst
C 34	16.4	86.3	139394	51	AC010046	AC010046 Drosophill
C 35	16.4	86.3	278077	73	AC010246	AC010246 Homo sapi
C 36	16.4	86.3	322667	34	AE003592	AE003592 Drosophill
C 37	16.4	84.2	2759	3	AF043942	AF043942 Bos tauru
C 38	16.4	84.2	2839	9	D86479	D86479 Homo sapien
C 39	16.4	84.2	153201	39	AC006454	AC006454 Homo sapi
C 40	16.4	84.2	153201	39	AC006454	AC006454 Homo sapi
C 41	15.8	83.2	1443	11	HS019601	U19601 Human (AML)
C 42	15.8	83.2	1710	9	D89788	D89788 Homo sapien
C 43	15.8	83.2	1710	9	D89789	D89789 Homo sapien
C 44	15.8	83.2	1710	9	D89790	D89790 Homo sapien
C 45	15.8	83.2	1728	3	AF047829	AF047829 Ovis arie

ALIGNMENTS

```
RESULT 1
LOCUS AF022953 1242 bp DNA PRI 30-OCT-1997
DEFINITION Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
ACCESSION AF022953
VERSION AF022953.1 GI:2570526
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1242)
Reihaus, E., Innis, M., MacIntyre, N. and Liggett, S.B.
Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL 93192047
MEDLINE 2 (bases 1 to 1242)
AUTHORS Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
A polymorphism of the human beta 2-adrenergic receptor within the
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL 94043092
MEDLINE 3 (bases 1 to 1242)
AUTHORS Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
Amino-terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL 94347707
MEDLINE Erratum: [[published erratum appears in Biochemistry 1994 Nov
29;33(47):14368]]
REMARK 4 (bases 1 to 1242)
REFERENCE Liggett, S.B. and Green, S.A.
Direct Submision
Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave ML670564, Cincinnati, OH 45267-0564, USA
JOURNAL
FEATURES
source
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADRB2"
1..1242
/gene="ADRB2"
/codon_start=1
/product="beta2-adrenergic receptor"
```

```
/protein_id="FAB82148.1"
/db_xref="GI:2570527"
/translation="MGCPNGSAFLIAPNGSHAPDHYNOQDEWVYVGMGIMSLIV
LAIVGNAVLTIRIAKFERLOQTVNFTISLACADLYMGLAVVPEAAHIILKMTFG
NFWCEFWISIDVLCVATASLETLICVAVNDYFPAITSFKQSLTKKAVIILMWIV
SGITSLPIOMHWYRATHOEAINCYANETCCDFEYNOAYIASIVSFYVPIVMEV
YSRVEQEARLOKIDKSGREYVONLSQEDDGRGHLRRSGFKLEKHEALKTIG
ITMGFTLCMLPEFIVYIHVDNLIRREYVILMWIGYVSGFNPILYCRSPDRI
APDEILCRRSSSKAAGNGYSSNGNGBDSGCHVEDEKKNLCELPOTEDPFVGHGS
TVPSDNIQSGRCSTNDSL"
46
/gene="ADRB2"
/feature="Arg16 to Gly polymorphism"
/replacement="a"
BASE COUNT 275 a 331 c 326 g 310 t
ORIGIN
Query Match 100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
RESULT 2
LOCUS AF022954 1242 bp DNA PRI 30-OCT-1997
DEFINITION Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
ACCESSION AF022954
VERSION AF022954.1 GI:2570528
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1242)
Reihaus, E., Innis, M., MacIntyre, N. and Liggett, S.B.
Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL 93192047
MEDLINE 2 (bases 1 to 1242)
AUTHORS Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
A polymorphism of the human beta 2-adrenergic receptor within the
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL 94043092
MEDLINE 3 (bases 1 to 1242)
AUTHORS Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
Amino-terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL 94347707
MEDLINE Erratum: [[published erratum appears in Biochemistry 1994 Nov
29;33(47):14368]]
REMARK 4 (bases 1 to 1242)
REFERENCE Liggett, S.B. and Green, S.A.
Direct Submision
Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave ML670564, Cincinnati, OH 45267-0564, USA
JOURNAL
FEATURES
source
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADRB2"
1..1242
/gene="ADRB2"
/codon_start=1
```

```
/product="beta2-adrenergic receptor"
/protein_id="AAB82149.1"
/db_xref="GI:2570529"
/translation="MGQPGNGSAFLIAPNRSHAPDHDTQGEDEVWVVGIMGLIVSLIV
LAIVGNVLVITAIKPERIQTIVNTVFITSLACADLVMGAVPFGAAHITLKKMTFG
NMCEFMISIDVLCYASIEFLCVIADVDFYFITSPPKQSLITFNKARVITLMMWIV
SGLTSEFLPIOMHMYRATHOEAINCYANETCCDFPTNOAVAIASSIVSPVPIVWIVFV
YSRVDEAKRQLOKIDKSGRPHVONLSQVEDGRGCHGLRRSSFCLEKHKALKTIG
IIMGFTTLCWLPFFIVNIVHYVIODLIRREVIITLNMIGYVNSGPNPLIYCSPDRI
AFQELICLRSSLKAVNGYSSNGMTGEOSGYHVEOEKENKLLCEDLPETEDFVGHOQ
TVPSNDISOGNCSTNDLSL"
79
/variation
/gene="ADRB2"
/feature="Gln27 to Glu polymorphism"
/replace="c"
BASE COUNT      276 a      330 c      326 g      310 t
ORIGIN
Query Match      100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 acatgacgatgccatgcc 19
Db      121 ACATGACGATGCCATGCC 103

RESULT 3
AF022955/c 1242 bp DNA PRI 30-OCT-1997
LOCUS Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION AF022955
ACCESSION AF022955.1 GI:2570530
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Reihnsaus,E., Innis,M., MacIntyre,N. and Liggett,S.B.
TITLE Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE 93192047
REFERENCE 2 (bases 1 to 1242)
AUTHORS Green,S.A., Cole,G., Jacinto,M., Innis,M. and Liggett,S.B.
TITLE A polymorphism of the human beta 2-adrenergic receptor within the
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL 94043092
MEDLINE
REFERENCE 3 (bases 1 to 1242)
AUTHORS Green,S.A., Turki,J., Innis,M. and Liggett,S.B.
TITLE Amino-terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
JOURNAL Biochemistry 33 (32), 9414-9419 (1994)
MEDLINE 94347707
REMARK Erratum:[published erratum appears in Biochemistry 1994 Nov
29;33(47):14368]]
4 (bases 1 to 1242)
REFERENCE 4 (bases 1 to 1242)
AUTHORS Liggett,S.B. and Green,S.A.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
source
1. .1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .1242
/feature="ADRB2"
CDS 1. .1242
/feature="ADRB2"
```

```
/codon_start=1
/product="beta2-adrenergic receptor"
/protein_id="AAB82150.1"
/db_xref="GI:2570531"
/translation="MGQPGNGSAFLIAPNRSHAPDHDTQGEDEVWVVGIMGLIVSLIV
LAIVGNVLVITAIKPERIQTIVNTVFITSLACADLVMGAVPFGAAHITLKKMTFG
NMCEFMISIDVLCYASIEFLCVIADVDFYFITSPPKQSLITFNKARVITLMMWIV
SGLTSEFLPIOMHMYRATHOEAINCYANETCCDFPTNOAVAIASSIVSPVPIVWIVFV
YSRVDEAKRQLOKIDKSGRPHVONLSQVEDGRGCHGLRRSSFCLEKHKALKTIG
IIMGFTTLCWLPFFIVNIVHYVIODLIRREVIITLNMIGYVNSGPNPLIYCSPDRI
AFQELICLRSSLKAVNGYSSNGMTGEOSGYHVEOEKENKLLCEDLPETEDFVGHOQ
TVPSNDISOGNCSTNDLSL"
100
/variation
/gene="ADRB2"
/feature="Val34 to Met polymorphism"
/replace="g"
BASE COUNT      277 a      331 c      324 g      310 t
ORIGIN
Query Match      100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 acatgacgatgccatgcc 19
Db      121 ACATGACGATGCCATGCC 103

RESULT 4
AF022956/c 1242 bp DNA PRI 30-OCT-1997
LOCUS Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION AF022956
ACCESSION AF022956.1 GI:2570532
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Reihnsaus,E., Innis,M., MacIntyre,N. and Liggett,S.B.
TITLE Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE 93192047
REFERENCE 2 (bases 1 to 1242)
AUTHORS Green,S.A., Cole,G., Jacinto,M., Innis,M. and Liggett,S.B.
TITLE A polymorphism of the human beta 2-adrenergic receptor within the
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL 94043092
MEDLINE
REFERENCE 3 (bases 1 to 1242)
AUTHORS Green,S.A., Turki,J., Innis,M. and Liggett,S.B.
TITLE Amino-terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
JOURNAL Biochemistry 33 (32), 9414-9419 (1994)
MEDLINE 94347707
REMARK Erratum:[published erratum appears in Biochemistry 1994 Nov
29;33(47):14368]]
4 (bases 1 to 1242)
REFERENCE 4 (bases 1 to 1242)
AUTHORS Liggett,S.B. and Green,S.A.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
source
1. .1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .1242
/feature="ADRB2"
CDS 1. .1242
/feature="ADRB2"
```

```
/gene="ADRB2"
/codon_start=1
/product="beta2-adrenergic receptor"
/protein_id="AB82151.1"
/db_xref="GI:2570533"
/translation="MGQPGNGSAFLAPNRSHAPDHDTQORDEVWVGMIYMSLIV
LAIVGNVITAIKPERLQVTNYETISLACADLVMGIAVVPFGAAHILMKMTG
NWCERFWSIDVLCVATSIETLCVIANDRFATSPFKQSILTKKARVITLMWIV
SGLISFLPIOMHWYRATHOEAIVANCYANETCCDEFTNOAIAISIVSFYPLIVMIV
YSRVFOEAKROLQIDKISGRFHVONLSOVEDGRTHGHLRSGKCLKEHKAUKTLTG
IIMGFTLCLMPFFIYIVIVIODNLIRKVEYILLMWIGVNSGFNPLIYCRSPDRRI
AFQELCLRRSSLKAYGNGYSSNGNTGSGYHVEDEKKNLCELDLPETDEPVGHOG
TVPSDNDISGGRNCSTNDLSL"
variation
491
/gene="ADRB2"
/note="Thr164 to Ile polymorphism"
/replace="C"
BASE COUNT      276 a      330 c      325 g      311 t
ORIGIN

Query Match      100.0%; Score 19; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1      acatgacgatgccatgcc 19
        |||||||
Db      121      ACATGACGATGCCCATGCC 103

RESULT      5
LOCUS      AF202305      1286 bp      DNA      PRI      14-DEC-1999
DEFINITION      Homo sapiens beta-2 adrenergic receptor gene, complete cds.
ACCESSION      AF202305
VERSION      AF202305.1      GI:6573152
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1286)
AUTHORS      Rupert,J.L. and Hochachka,P.W.
TITLES      Beta-2-adrenergic receptor allele frequencies in two native
            American populations
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1286)
AUTHORS      Rupert,J.L. and Hochachka,P.W.
TITLES      Direct Submission
JOURNAL      Submitted (04-NOV-1999) Zoology, University of British Columbia,
            6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
            Location/Qualifiers
FEATURES
    source
        1..1286
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="5"
            /map="5q31-q33"
            /note="Isolated from Quechua speaking Native American"
            <35..>1276
            /product="beta-2 adrenergic receptor"
            35..1276
            /codon_start=1
            /product="beta-2 adrenergic receptor"
            /protein_id="AA017569.1"
            /db_xref="GI:6573153"
            /translation="MGQPGNGSAFLAPNRSHAPDHDTQORDEVWVGMIYMSLIV
            LAIVGNVITAIKPERLQVTNYETISLACADLVMGIAVVPFGAAHILMKMTG
            NWCERFWSIDVLCVATSIETLCVIANDRFATSPFKQSILTKKARVITLMWIV
            SGLISFLPIOMHWYRATHOEAIVANCYANETCCDEFTNOAIAISIVSFYPLIVMIV
            YSRVFOEAKROLQIDKISGRFHVONLSOVEDGRTHGHLRSGKCLKEHKAUKTLTG
            IIMGFTLCLMPFFIYIVIVIODNLIRKVEYILLMWIGVNSGFNPLIYCRSPDRRI
            AFQELCLRRSSLKAYGNGYSSNGNTGSGYHVEDEKKNLCELDLPETDEPVGHOG
            TVPSDNDISGGRNCSTNDLSL"
CDS
    mRNA
    CDS
```

```
BASE COUNT      282 a      347 c      334 g      320 t      3 others
ORIGIN

Query Match      100.0%; Score 19; DB 39; Length 1286;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1      acatgacgatgccatgcc 19
        |||||||
Db      155      ACATGACGATGCCCATGCC 137

RESULT      6
LOCUS      AF169225      1290 bp      DNA      PRI      10-AUG-1999
DEFINITION      Homo sapiens beta-2-adrenergic receptor gene, complete cds.
ACCESSION      AF169225
VERSION      AF169225.1      GI:5714687
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 1290)
AUTHORS      Rupert,J.R. and Hochachka,P.W.
TITLES      Beta-2-adrenergic receptor allele frequencies in two native
            American populations
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1290)
AUTHORS      Rupert,J.R. and Hochachka,P.W.
TITLES      Direct Submission
JOURNAL      Submitted (14-JUL-1999) Zoology, University of British Columbia,
            6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
            Location/Qualifiers
FEATURES
    source
        1..1290
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="5"
            /map="5q31-q33"
            /cell_type="lymphocyte"
            /tissue_type="blood"
            /note="Isolated from a Quechua speaking native American
            heterozygous for a known C/T mutation"
            17..1258
            /codon_start=1
            /product="beta-2-adrenergic receptor"
            /protein_id="AA048036.1"
            /db_xref="GI:5714688"
            /translation="MGQPGNGSAFLAPNRSHAPDHDTQORDEVWVGMIYMSLIV
            LAIVGNVITAIKPERLQVTNYETISLACADLVMGIAVVPFGAAHILMKMTG
            NWCERFWSIDVLCVATSIETLCVIANDRFATSPFKQSILTKKARVITLMWIV
            SGLISFLPIOMHWYRATHOEAIVANCYANETCCDEFTNOAIAISIVSFYPLIVMIV
            YSRVFOEAKROLQIDKISGRFHVONLSOVEDGRTHGHLRSGKCLKEHKAUKTLTG
            IIMGFTLCLMPFFIYIVIVIODNLIRKVEYILLMWIGVNSGFNPLIYCRSPDRRI
            AFQELCLRRSSLKAYGNGYSSNGNTGSGYHVEDEKKNLCELDLPETDEPVGHOG
            TVPSDNDISGGRNCSTNDLSL"
            491
            /replace="C/T"
variation
    BASE COUNT      287 a      349 c      331 g      322 t      1 others
ORIGIN

Query Match      100.0%; Score 19; DB 39; Length 1290;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1      acatgacgatgccatgcc 19
        |||||||
Db      137      ACATGACGATGCCCATGCC 119

RESULT      7
```

AF203386/c
 LOCUS AF203386 1290 bp DNA PRI 28-DEC-1999
 DEFINITION Homo sapiens beta-2 adrenergic receptor (ADRB2) gene, complete cds.
 ACCESSION AF203386
 VERSION AF203386.1 GI:6636495
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1290)
 AUTHORS Rupert, J.L. and Hochachka, P.W.
 TITLE Beta-2 adrenergic receptor allele frequencies in two Native American populations
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1290)
 AUTHORS Rupert, J.L. and Hochachka, P.W.
 TITLE Direct Submission
 JOURNAL Submitted (09-NOV-1999) Zoology, University of British Columbia,
 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
 FEATURES
 source
 1..1290
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q32-q34"
 /sex="female"
 /note="Isolated from a Quechuan-speaking Native American"
 26..1267
 /gene="ADRB2"
 26..1267
 /gene="ADRB2"
 /note="catecholamine receptor"
 /product="beta-2 adrenergic receptor"
 /protein_id="AAF20199.1"
 /db_xref="GI:6636496"
 /translation="MGPGNGSAFLAPNGSHAPDHDTQQRDEWVYVGMGIVMSLIY
 LAIVFGNVLVITAIKFERLQVTNFITSLACADLVGLAVPGASHIILKMTFG
 NMCCEPMTSIDVLCVASTIETLCVIAVDYFAITSPKQSLITNKARVILMWIV
 SGLTSPFLPIOMHMYRATHQEAINECYANETCCDFITNOAVIAISSIVSEYPLVWVY
 YSRVFQVAKROLOKIDRSRGRHQAONLSQVEDDGRSGHGRHSRSCFLKEHAKLTIG
 IIMGTFTLCMLPEFTVNIYVHVDNLIPREYVILNWIGYVNSAPNPLIYCSPPERI
 AFQELICLRSSLSKAYNGYSSNGNTGDSGTHVDEKKNLCELDPEDEPVGHGQ
 TVPSNDISDQGNCSITNDSL"
 BASE COUNT 288 a 345 c 333 g 324 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 39; Length 1290;
 Best local Similarity 100.0%; Pred. No. 51;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 acatgcagatgccatgcc 19
 ||||||||||||||||
 Db 146 ACATGACATGCCATGCC 128
 RESULT 8
 CFU73206 1298 bp DNA MM 31-DEC-1997
 LOCUS CFU73206/c
 DEFINITION Canis familiaris beta2 adrenergic receptor (dogbeta2) gene,
 complete cds.
 ACCESSION U73206
 VERSION U73206.1 GI:1657819
 KEYWORDS
 SOURCE dog.
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Carnivora; Filippidae; Canidae; Canis.
 REFERENCE 1 (bases 1 to 1298)
 AUTHORS Huang, R.R., Rapoport, D., Schaeffer, M.T., Cascleri, M.A. and
 Fong, T.M.

TITLE Molecular cloning of the dog beta 1 and beta 2 adrenergic receptors
 JOURNAL J. Recept. Signal Transduct. Res. 17 (4), 599-607 (1997)
 MEDLINE 97364078
 REFERENCE 2 (bases 1 to 1298)
 AUTHORS Huang, R.-R.C., Rapoport, D., Schaeffer, M.-T., Cascleri, M.A. and
 Fong, T.M.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-1996) R80W-213, Merck, 126 E. Lincoln Ave,
 Rahway, NJ 07065, USA
 FEATURES
 source
 1..1298
 /organism="Canis familiaris"
 /db_xref="taxon:9615"
 14..1261
 /gene="dogbeta2"
 14..1261
 /gene="dogbeta2"
 /note="G protein coupled receptor"
 /product="beta2 adrenergic receptor"
 /protein_id="BAB93647.1"
 /db_xref="GI:1657820"
 /translation="MGOPANRSVFLAPNGSHAPDQDSQERSEAMVYVGMGIVMSLIY
 LAIVFGNVLVITAIKFERLQVTNFITSLACADLVGLAVPGASHIILKMTFG
 NMCCEPMTSIDVLCVASTIETLCVIAVDYFAITSPKQSLITNKARVILMWIV
 SGLTSPFLPIOMHMYRATHQEAINECYANETCCDFITNOAVIAISSIVSEYPLVWVY
 YSRVFQVAKROLOKIDRSRGRHQAONLSQVEDDGRSGHGRHSRSCFLKEHAKLTIG
 IIMGTFTLCMLPEFTVNIYVHVDNLIPREYVILNWIGYVNSAPNPLIYCSPPERI
 AFQELICLRSSLSKAYNGYSSNGNTGDSGTHVDEKKNLCELDPEDEPVGHGQ
 OCTVPSDVSQGNCSITNDSL"
 BASE COUNT 260 a 396 c 359 g 283 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 3; Length 1298;
 Best local Similarity 100.0%; Pred. No. 51;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 acatgcagatgccatgcc 19
 ||||||||||||||||
 Db 134 ACATGACATGCCATGCC 116
 RESULT 9
 MACB2AR/c
 LOCUS MACB2AR/c
 DEFINITION Macaca mulatta beta-2 adrenergic receptor (B2AR) mRNA, complete
 cds.
 ACCESSION L38905
 VERSION L38905.1 GI:1004338
 KEYWORDS beta-2 adrenergic receptor.
 SOURCE Macaca mulatta cDNA to mRNA.
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecoinae;
 Macaca.
 REFERENCE 1 (bases 1 to 1320)
 AUTHORS Amend, A.M. and Guan, X.M.
 TITLE Cloning, sequencing, and expression of the rhesus monkey beta 2
 adrenergic receptor
 JOURNAL DNA Cell Biol. 14 (9), 753-757 (1995)
 MEDLINE 95398843
 FEATURES
 source
 1..1320
 /organism="Macaca mulatta"
 /db_xref="taxon:9544"
 /note="genomic DNA for this receptor sequenced and found
 to be identical with the cDNA"
 1..40
 /gene="B2AR"
 1..1320
 /gene="B2AR"
 1..1320

5'UTR
 gene
 mRNA

```

CDS
    /gene="B2AR"
    41..1288
    /gene="B2AR"
    /codon_start=1
    /product="beta-2 adrenergic receptor"
    /protein_id="AAC41914.1"
    /db_xref="GI:1004339"
    /translation="MGDPGSAFLAPNGSHAPDHDVTQERDEAVVVGMIYMSLIV
    LAIVGNAVITAIARPERLOTVTNFTISLACADLVKGLAVPFGASHILKKMTFG
    NFWCEWTSIDVLCVTAISIEITLCVIAVDYFAITSPFKYQSLLTKKARVILMWIV
    SGLTSLPIQMHWYRATHOEAINCYAKERCDFPTNOAIAISSIVSFYLVVWV
    YSRVFOAKROLOKIDSRGFRHONLSQVEDGRSGHGRSSKFCLEKHAALTKIG
    IIMGTFTLWLPFTYINIVAVIODNLIPKEVYLLMWGVYNSAFNPFLYCRSPDRI
    AFQELCLRRSLKATGNGYSNNSNGTGEOSGYHLEQEKENLLCEDLPGEDEVGH
    OGTVPSDSDVSDGRCSTNDLSL"

3'UTR
    1289..1320
    /gene="B2AR"
    1289..1320

BASE COUNT      289 a      365 c      343 g      323 t

ORIGIN
    Query Match      100.0%; Score 19; DB 11; Length 1320;
    Best Local Similarity 100.0%; Pred. No. 51;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 acatgacgatgccatgcc 19
    |||||
    Db 161 ACATGACGATGCCATGCC 143

RESULT 10
AF192345 1536 bp DNA MAM 27-OCT-1999
LOCUS AF192345
DEFINITION Felis catus beta-2 adrenergic receptor gene, complete cds.
ACCESSION AF192345
VERSION AF192345.1 GI:6120128

KEYWORDS
    cat.
    SOURCE
    ORGANISM
        Felis catus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
        Eutheria; Carnivora; Fissipedia; Felidae; Felis.
    REFERENCE
        1 (bases 1 to 1536)
        Cully,D.F., Tremml,G. and Zachwieja,S.
        Felis domesticus beta adrenergic receptor subtype 2
        Unpublished
        2 (bases 1 to 1536)
        Cully,D.F., Tremml,G. and Zachwieja,S.
        Direct Submission
        Submitted (05-OCT-1999) MRL, Merck & Co., Box 2000 R80N-C42,
        Rahway, NJ 07065, USA
    JOURNAL
        Location/Qualifiers
        1..1536
        /organism="Felis catus"
        /db_xref="taxon:9685"
        /note="synonym: Felis domesticus"
        <6..>1262
        /product="beta-2 adrenergic receptor"
        6..1262
        /note="G-protein coupled receptor; beta adrenergic
        receptor subtype 2"
        /codon_start=1
        /product="beta-2 adrenergic receptor"
        /protein_id="AAF04304.1"
        /db_xref="GI:6120129"
        /translation="MGDPGSAFLAPNGSHAPDHDVTQERDEAVVVGMIYMSLIV
        LAIVGNAVITAIARPERLOTVTNFTISLACADLVKGLAVPFGASHILKKMTFG
        NFWCEWTSIDVLCVTAISIEITLCVIAVDYFAITSPFKYQSLLTKKARVILMWIV
        SGLTSLPIQMHWYRATHOEAINCYAKERCDFPTNOAIAISSIVSFYLVVWV
        YSRVFOAKROLOKIDSRGFRHONLSQVEDGRSGHGRSSKFCLEKHAALTKIG
        IIMGTFTLWLPFTYINIVAVIODNLIPKEVYLLMWGVYNSAFNPFLYCRSPDRI
        AFQELCLRRSLKATGNGYSNNSNGTGEOSGYHLEQEKENLLCEDLPGEDEVGH
        OGTVPSDSDVSDGRCSTNDLSL"

BASE COUNT      342 a      424 c      398 g      372 t

```

```

ORIGIN
    Query Match      100.0%; Score 19; DB 3; Length 1536;
    Best Local Similarity 100.0%; Pred. No. 51;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 acatgacgatgccatgcc 19
    |||||
    Db 126 ACATGACGATGCCATGCC 108

RESULT 11
CFB2AR/c 1948 bp mRNA MAM 27-JAN-1997
LOCUS CFB2AR
DEFINITION C.familiaris mRNA for beta2-adrenergic receptor.
ACCESSION X94608
VERSION X94608.1 GI:1359588
KEYWORDS
    beta-2 adrenergic receptor.
    dog.
    SOURCE
    ORGANISM
        Canis familiaris
        Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
        Carnivora; Fissipedia; Canidae; Canis.
    REFERENCE
        1 (bases 1 to 1948)
        Emla,C.W., Kuhl,J., Hirschman,C.A. and Levine,M.A.
        Rapid communication: cloning and sequencing of a canine beta
        2-adrenergic receptor cDNA
        J Anim. Sci. 74 (9), 2285 (1996)
    JOURNAL
        97034778
        2 (bases 1 to 1948)
        Emla,C.W.
        Direct Submission
        Submitted (29-DEC-1995) C.W. Emla, Johns Hopkins School of
        Medicine, Dept of Anesthesiology, John Hopkins Hospital, Meyer
        297a, Baltimore Maryland 21287, USA
    JOURNAL
        Location/Qualifiers
        1..1948
        /organism="Canis familiaris"
        /db_xref="taxon:9615"
        /tissue_type="cardiac muscle"
        151..1398
        /codon_start=1
        /product="beta2-adrenergic receptor"
        /protein_id="CAA64316.1"
        /db_xref="GI:1359589"
        /db_xref="SWISS-PROT:P54833"
        /translation="MGDPANRSVFLAPNGSHAPDGDGQSESEAVVVGMIYMSLIV
        LAIVGNAVITAIARPERLOTVTNFTISLACADLVKGLAVPFGASHILKKMTFG
        NFWCEWTSIDVLCVTAISIEITLCVIAVDYFAITSPFKYQSLLTKKARVILMWIV
        SGLTSLPIQMHWYRATHOEAINCYAKERCDFPTNOAIAISSIVSFYLVVWV
        YSRVFOAKROLOKIDSRGFRHONLSQVEDGRSGHGRSSKFCLEKHAALTKIG
        IIMGTFTLWLPFTYINIVAVIODNLIPKEVYLLMWGVYNSAFNPFLYCRSPDRI
        AFQELCLRRSLKATGNGYSNNSNSDIDAGHSCHLQGEKDSBLCEDPPGIEDR
        OGTVPSDSDVSDGRCSTNDLSL"

BASE COUNT      395 a      539 c      540 g      474 t

ORIGIN
    Query Match      100.0%; Score 19; DB 3; Length 1948;
    Best Local Similarity 100.0%; Pred. No. 51;
    Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 acatgacgatgccatgcc 19
    |||||
    Db 271 ACATGACGATGCCATGCC 253

RESULT 12
HSB2AR/c 1970 bp mRNA PRI 12-SEP-1993
LOCUS HSB2AR
DEFINITION Human mRNA for brain beta-adrenergic receptor.
ACCESSION X04827
VERSION X04827.1 GI:29372

```

KEYWORDS beta-adrenergic receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1970)
AUTHORS Chung, F.Z., Lentes, K.U., Gocayne, J., Fitzgerald, M., Robinson, D.,
Kerlavage, A.R., Fraser, C.M. and Venter, J.C.
TITLE Cloning and sequence analysis of the human brain beta-adrenergic
receptor. Evolutionary relationship to rodent and avian
beta-receptors and porcine muscarinic receptors
JOURNAL FEBS Lett. 211 (2), 200-206 (1987)
MEDLINE 87105974
REFERENCE 2 (bases 1 to 1970)
AUTHORS Kerlavage, A.R.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1987) to the EMBL/Genbank/DBJ databases
COMMENT Substantial corrections are reported in [2]
Data kindly reviewed (22-SEP-1987) by Kerlavage A.R.

FEATURES
Source
1. .1970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="neonatal human brain stem"
178. .1419
/note="beta-adrenergic receptor (AA 1-413)"
/codon_start=1
/protein_id="CAA28511.1"
/db_xref="GI:29373"
/db_xref="SWISS-PROT:P07550"
/translation="MGORGNSAFLLAPNSHAPDHYTOERDEVVYVGMGLVMSLIY
LAIVGNAVITTAIKKPERLOTVTNYFTTSLACADLVNGLAVYVPGAAHIILKKMTFC
NFWCEFWTSIDLVCYTAIEITLCVAVDRFAITSPRYOSLITNKARVILIMWIV
SGLTSLPDIOMHWKATQHEAIVNCYANETCDPTNQAVYALASSIVSFYVPLVYVWV
YSRVEQBAKROLOKIDKSGEGRHYONLSQVEDQRTGGLRRSSFCLEKHKALKTG
IMGFETLCMLPEFIVNIYVHIDNLIRKEYIILNMIGYVNSGPNPLCYGSPERI
AFOLLCILRRSLKAYGNSNGTGTGSGYHVQEKENKLCEDLPGTEDFVGHOG
TVPSNDISQGRNCSTNDSL"

CDS
misc_feature
794. .799
/note="pot. glucocorticoid-responsive element"
misc_feature
965. .970
/note="pot. glucocorticoid-responsive element"
misc_feature
1459. .1464
/note="pot. glucocorticoid-responsive element"
misc_feature
1491. .1496
/note="pot. glucocorticoid-responsive element"
misc_feature
1502. .1507
/note="pot. glucocorticoid-responsive element"
misc_feature
1952. .1957
/note="pot. glucocorticoid-responsive element"
misc_feature
1970
/note="pot. glucocorticoid-responsive element"
BASE COUNT 459 a 508 c 482 g 521 t
ORIGIN

Query Match 100.0%; Score 19; DB 9; Length 1970;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 acatgacgatgccatgcc 19
|||||
Db 298 ACATGACGATGCCATGCC 280

RESULT 13
LOCUS HSBAR 2305 bp DNA PRI 12-SEP-1993
DEFINITION Human gene for beta-adrenergic receptor (beta-2 subtype).
ACCESSION Y00106
VERSION Y00106.1 GI:29370
KEYWORDS beta-adrenergic receptor.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2305)
AUTHORS Schofield, P.R., Rhee, L.M. and Peralta, E.G.
TITLE Primary structure of the human beta-adrenergic receptor gene
JOURNAL Nucleic Acids Res. 15 (8), 3636 (1987)
MEDLINE 87203400
REFERENCE 2 (bases 1 to 2305)
AUTHORS Schofield, P.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1987) to the EMBL/Genbank/DBJ databases
COMMENT

FEATURES
Source
1. .2305
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="Maniacis human"
794. .2035
/note="beta-adrenergic receptor (AA 1 - 413)"
/codon_start=1
/protein_id="CAA68289.1"
/db_xref="GI:29371"
/db_xref="SWISS-PROT:P07550"
/translation="MGORGNSAFLLAPNSHAPDHYTOERDEVVYVGMGLVMSLIY
LAIVGNAVITTAIKKPERLOTVTNYFTTSLACADLVNGLAVYVPGAAHIILKKMTFC
NFWCEFWTSIDLVCYTAIEITLCVAVDRFAITSPRYOSLITNKARVILIMWIV
SGLTSLPDIOMHWKATQHEAIVNCYANETCDPTNQAVYALASSIVSFYVPLVYVWV
YSRVEQBAKROLOKIDKSGEGRHYONLSQVEDQRTGGLRRSSFCLEKHKALKTG
IMGFETLCMLPEFIVNIYVHIDNLIRKEYIILNMIGYVNSGPNPLCYGSPERI
AFOLLCILRRSLKAYGNSNGTGTGSGYHVQEKENKLCEDLPGTEDFVGHOG
TVPSNDISQGRNCSTNDSL"

CDS
misc_feature
809. .817
/note="N-linked glycosylation site"
misc_feature
836. .844
/note="N-linked glycosylation site"
misc_feature
896. .967
/note="membrane spanning domain I"
misc_feature
1007. .1078
/note="membrane spanning domain II"
misc_feature
1114. .1180
/note="membrane spanning domain III"
misc_feature
1247. .1315
/note="membrane spanning domain IV"
misc_feature
1383. .1450
/note="membrane spanning domain V"
misc_feature
1616. .1687
/note="membrane spanning domain VI"
misc_feature
1712. .1774
/note="membrane spanning domain VII"
BASE COUNT 495 a 616 c 649 g 545 t
ORIGIN

Query Match 100.0%; Score 19; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 acatgacgatgccatgcc 19
|||||
Db 914 ACATGACGATGCCATGCC 896

RESULT 14
LOCUS A65720 2679 bp DNA PAT 29-MAR-1999
DEFINITION Sequence 1 from Patent WO9735973.
ACCESSION A65720
VERSION A65720.1 GI:4531340
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 2679)
 AUTHORS Lensen,G., Pietri-Rouxel,F., Drumare, Marie-Francoise and
 Scroberg,A.D.
 TITLE CANINE beta 2- AND beta 3-ADRENERGIC RECEPTORS AND USE THEREOF
 JOURNAL Patent: WO 9735973-A 02-Oct-1997;
 VERIGEN (FR)
 COMMENT Other publication FR 2746813 19971003.
 FEATURES Location/Qualifiers
 source 1..2679
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 577 a 736 c 724 g 642 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 5; Length 2679;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 acatgacgatgccatgcc 19
 ||||||||||||||||
 Db 289 ACATGACGATGCCATGCC 271
 RESULT 15
 HUMADRB/C 3451 bp mRNA PRI 13-FEB-1996
 LOCUS Human beta-2-adrenergic receptor mRNA, complete cds.
 ACCESSION M15169 J02728 M16106
 VERSION M15169.1 GI:178201
 KEYWORDS adrenergic receptor.
 SOURCE Homo sapiens (clone: pTF.) (tissue library: Ewan Sadler) placenta
 CDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 3451)
 AUTHORS Koblika,B.K., Friele,T., Dohman,H.G., Bolanowski,M.A.,
 Dixon,R.A., Keller,P., Caron,M.G. and Lefkowitz,R.J.
 TITLE Delineation of the intronless nature of the genes for the human and
 hamster beta 2-adrenergic receptor and their putative promoter
 regions
 JOURNAL J. Biol. Chem. 262 (15), 7321-7327 (1987)
 MEDLINE 87222338
 REFERENCE 2 (bases 1399 to 1985)
 AUTHORS Koblika,B.K., Dixon,R.A., Friele,T., Dohman,H.G.,
 Bolanowski,M.A., Sigal,I.S., Yang-Feng,T.L., Francke,U., Caron,M.G.
 and Lefkowitz,R.J.
 TITLE cDNA for the human beta 2-adrenergic receptor: a protein with
 multiple membrane-spanning domains and encoded by a gene whose
 chromosomal location is shared with that of the receptor for
 platelet-derived growth factor
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (1), 46-50 (1987)
 MEDLINE 87092393
 FEATURES Location/Qualifiers
 source 1..3451
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="pTF."
 /tissue_type="Placenta"
 /tissue_id="Ewan Sadler"
 /map="5q31-q32"
 1369..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.); G00-120-541"
 1369..3383
 /gene="ADRB2"
 1376..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.); G00-120-541"
 1379..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.); G00-120-541"

MRNA 1388..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.); G00-120-541"
 CDS 1487..1546
 /gene="ADRB2"
 /note="putative"
 /codon_start=1
 /db_xref="GDB:G00-120-541"
 /protein_id="AA88014.1"
 /db_xref="GI:560761"
 /translation="MRLGVNSRPAERPGSAR"
 1588..2829
 /gene="ADRB2"
 /codon_start=1
 /db_xref="GDB:G00-120-541"
 /product="beta-2 adrenergic receptor"
 /protein_id="AA88015.1"
 /db_xref="GI:178202"
 /translation="MGOPNGSAFLAPNRSHPADHDVTOQRDEWYVYGNIWSTLY
 LAIVGNVITAIKFERIQVTNYPFTSLACADLVGLAVVPGAAHILMKWITG
 NFWCEFTSIDVLTASIEITLCVAVDRYFAITSPEKYOSLTKRNARVILMWIV
 SGITSEFLIDHWRATHTHDAINCYANETCCDFTNQAIASSIVSFYVLTWIV
 YSRVQEAQRLOKIDKSEGRPHONTSOVRDQRTGHLGRSSKFCLEKALKTLG
 IIMGFTLCWLPFTLVNTVHYIQDLIRKEYITILNIGYVNSGFNPPLYCRSPDRI
 AFQELICLRSSLSKAVNGYSSNGVTGQSGIHVEQEKENKLCEDLPGTEDFVGHGQ
 TVPSNDISQGRNCSTNDSL"

BASE COUNT 790 a 873 c 895 g 893 t
 ORIGIN
 Query Match 100.0%; Score 19; DB 10; Length 3451;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 acatgacgatgccatgcc 19
 ||||||||||||||||
 Db 1708 ACATGACGATGCCATGCC 1690

Search completed: September 12, 2000, 23:02:50
 Job time: 3962 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:07:17 : Search time 148.16 seconds
(without alignments)
32.085 Million cell updates/sec

Title: US-09-542-718-3

Perfect score: 19
Sequence: 1 acatgacgatgccatgcc 19

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	187	1	Q13384
2	19	100.0	191	1	ZC2720 oligonucleo
3	19	100.0	1287	1	ZC2719 oligonucleo
4	19	100.0	1400	1	Human beta2-adrene
5	19	100.0	1999	1	Beta-2 adrenalin r
6	19	100.0	2679	1	Canine beta-2 adre
7	19	100.0	3451	1	Human beta-2 adre
8	17.4	91.6	432	1	EST clone DU675. N
9	16	84.2	2795	1	Human E2A-binding
10	16	84.2	3854	1	Human E2A-binding
11	15.4	81.1	4131	1	Acetobacter cdg3 o
12	14.8	77.9	453	1	Pathogenesis-relat
13	14.8	77.9	480	1	Pathogenesis-relat
14	14.8	77.9	730	1	GGF gene coding se
15	14.8	77.9	730	1	GGF gene coding se
16	14.8	77.9	730	1	Human glial cell
17	14.8	77.9	730	1	Human glial growth
18	14.8	77.9	730	1	Human glial growth
19	14.8	77.9	730	1	Human glial growth
20	14.8	77.9	1063	1	Human NDF-beta1a c
21	14.8	77.9	1164	1	Human NDF-alpha1a
22	14.8	77.9	1651	1	Complete gene sequ
23	14.8	77.9	1807	1	Human NDF-alpha2b
24	14.8	77.9	1823	1	Human NDF-alpha2b
25	14.8	77.9	1873	1	DNA sequence of Lu
26	14.8	77.9	2010	1	Human NDF-beta1a c
27	14.8	77.9	2029	1	Human heregulin al
28	14.8	77.9	2164	1	Human heregulin be
29	14.8	77.9	2164	1	Human breast cance
30	14.8	77.9	2199	1	Human breast cance
31	14.8	77.9	2199	1	Human breast cance
32	14.8	77.9	2335	1	Full cDNA sequence
33	14.8	77.9	2356	1	Human PRONDF-alpha
					Rat NDF clone 38 D

34	14.8	77.9	2430	1	080232	Rat NDF clone 42A
35	14.8	77.9	2531	1	080227	Rat NDF clone 20 D
36	14.8	77.9	2743	1	080231	Rat NDF clone 41 D
37	14.8	77.9	2914	1	080226	Rat NDF clone 19 D
38	14.8	77.9	3060	1	T62781	Urea transporter P
39	14.8	77.9	3161	1	080233	Rat NDF clone 42B
40	14.8	77.9	3344	1	080228	Rat NDF clone 22 D
41	14.8	77.9	4018	1	063879	Polyhydroxyalkanoa
42	14.4	75.8	430	1	V87528	EST clone DN381. N
43	14.4	75.8	1898	1	054624	Mammalian vesicle
44	14.4	75.8	1898	1	T95867	cDNA for human syn
45	14.4	75.8	2228	1	034554	cDNA encoding rp40

ALIGNMENTS

RESULT 1	1	100.0%	Score 19:	DB 1:	Length 187:
ID	013384	standard; DNA: 187 BP.			
AC	013384;				
DT	07-NOV-1991 (first entry)				
DE	ZC2720 oligonucleotide (PHRS17).				
KM	AR: adrenergic receptor; G-protein: ss.				
PN	WO9112273-A.				
PD	22-AUG-1991.				
PF	08-FEB-1991: U00909.				
PR	08-FEB-1990: US-478100.				
PA	(ZYMO-) ZYMOGENETICS INC.				
PI	Slodziewsk AZ, Sheppard PO;				
DR	WPI: 91-267100/36.				
PT	Producing hybrid G-protein coupled receptors - mammalian DNA				
PT	sequences encoding receptors having at least one domain replaced				
PS	by corresp. yeast domain used to transform yeast host cells				
CC	Disclosure: Page 27: 65pp: English.				
CC	Oligonucleotides ZC2719 and ZC2720 (Q13383-84) were designed to				
CC	encode a 5' end by an EcoRI adhesive and followed by the extracellular				
CC	N-terminal domain of the STE2 gene prod. contg. nucleotides 1-147 of the				
CC	yeast G protein-coupled protein gene joined to nucleotides 103-116 of				
CC	the human G protein-coupled protein gene (see Q13381).				
CC	The oligonucleotides, the beta2AR fragment and the pmv1 vector				
CC	fragment were joined in a four-part ligation, to construct PHRS16.				
CC	From PHRS16, construct PHRS17 is produced encoding the human beta2-				
CC	adrenergic-STE2 hybrid receptor.				
CC	See also Q13381 and Q13383-83.				
SQ	Sequence 187 BP; 58 A; 36 C; 43 G; 50 T;				
Query Match	100.0%	Score 19:	DB 1:	Length 187:	
Best Local Similarity	100.0%	Pred. NO. 0.47:			
Matches 19; Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;				
QY	1	acatgacgatgccatgcc 19			
DB	16	ACATGACGATGCCATGCC 34			
RESULT 2	2	100.0%	Score 19:	DB 1:	Length 187:
ID	013383/C	standard; DNA: 191 BP.			
AC	013383;				
DT	07-NOV-1991 (first entry)				
DE	ZC2719 oligonucleotide (PHRS17).				
KM	AR: adrenergic receptor; G-protein: ss.				
PN	WO9112273-A.				
PD	22-AUG-1991.				
PF	08-FEB-1991: U00909.				
PR	08-FEB-1990: US-478100.				
PA	(ZYMO-) ZYMOGENETICS INC.				
PI	Slodziewsk AZ, Sheppard PO;				
DR	WPI: 91-267100/36.				
PT	Producing hybrid G-protein coupled receptors - mammalian DNA				
PT	sequences encoding receptors having at least one domain replaced				

PT by corresp. yeast domain used to transform yeast host cells
 PS Disclosure: Page 27: 65pp. English.
 CC Oligonucleotides ZC2719 and ZC2720 (Q13383-84) were designed to
 CC encode a 5' end by an EcoRI adhesive end followed by the extracellular
 CC N-terminal domain of the STE2 gene prod. contg. nucleotides 1-147 of the
 CC yeast G protein-coupled protein gene joined to nucleotides 103-136 of
 CC the human G protein-coupled protein gene (see Q13381).
 CC The oligonucleotides, the beta2AR fragment and the pMVR1 vector
 CC fragment were joined in a four-part ligation, to construct pHR516.
 CC From pHR516, construct pHR517 is produced encoding the human beta2-
 CC adrenergic-STE2 hybrid receptor.
 CC See also Q13381 and Q13383-83.
 SQ Sequence 191 BP; 52 A; 43 C; 36 G; 60 T;

Query Match 100.0%; Score 19; DB 1; Length 191;
 Best Local Similarity 100.0%; Pred. No. 0.47;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acatgacgatgccatgcc 19
 Db 176 ACATGACGATGCCATGCC 158

RESULT 3
 Q13381/c
 ID Q13381 standard; DNA; 1287 BP.
 AC Q13381;
 DT 07-NOV-1991 (first entry)
 DE Human beta2-adrenergic-STE2 hybrid receptor (1).
 KW AR: adrenergic receptor; G-protein; ligand; screening; internal;
 KW extracellular; N-terminal; C-terminal; effector; domain; ss.
 FH Key Location/Qualifiers
 FT misc_rna
 FT 1..147
 FT /*tag= a
 FT /note= "nucleotides 1-147 of yeast G
 FT protein-coupled receptor."
 FT 148..1287
 FT /*tag= b
 FT /note= "nucleotides 103-1242 of human G protein-
 FT coupled receptor"

misc_rna
 FT 148..1287
 FT /*tag= b
 FT /note= "nucleotides 103-1242 of human G protein-
 FT coupled receptor"

W09112273-A.
 PN 22-AUG-1991.
 PR 08-FEB-1991; U00909.
 PR 08-FEB-1991; US-478100.
 PA (ZYMO-) ZYMOGENETICS INC.
 PI Sledziewsk AZ, Sheppard PO;
 DR WPI: 91-267100/36.
 DR P-PSDB: R13729.
 PT Producing hybrid G-protein coupled receptors - mammalian DNA
 PT sequences encoding receptors having at least one domain replaced
 PT by corresp. yeast domain used to transform yeast host cells
 PS Disclosure: Page 37-38; 65pp; English.
 CC The hybrid beta2AR-STE2 receptor was constructed by replacing the DNA
 CC sequence encoding the extracellular N-terminal domain with a DNA
 CC sequence encoding the extracellular N-terminal domain of the
 CC Saccharomyces cerevisiae STE2 gene product.
 CC Construct pHR517 comprises this sequence together with a TPII
 CC transcriptional promoter and terminator. The construct is used to
 CC transform yeast cells. For the detection of the presence of a ligand in
 CC a test sample, the culture of the yeast cells may be exposed to the test
 CC sample, under conditions allowing binding of the receptor and ligand.
 CC Response to the ligands can be easily monitored. The method of
 CC screening for potential G protein-coupled receptor ligands is less
 CC expensive and labour intensive than previous methods and does not
 CC necessitate the isolation of membrane fragments from responsive
 CC tissues or cell lines.
 CC Other exemplary DNA sequences encoding hybrid G protein-coupled
 CC receptors are as follows: (1) the C-terminal internal and C-terminal
 CC domain is replaced; (2) the N-terminal extracellular and C-terminal
 CC internal effector domains are replaced; and (3) the N-terminal
 CC extracellular domain, the third internal effector domain and C-
 CC terminal internal effector domain are replaced; all with the

CC corresp. domains of the S. cerevisiae STE2 gene.
 CC See also Q13383-83.
 SQ Sequence 1287 BP; 294 A; 333 C; 314 G; 346 T;

Query Match 100.0%; Score 19; DB 1; Length 1287;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acatgacgatgccatgcc 19
 Db 166 ACATGACGATGCCATGCC 148

RESULT 4
 T93249/c
 ID T93249 standard; cDNA to mRNA; 1400 BP.
 AC T93249;
 DT 20-APR-1998 (first entry)
 DE Beta-2 adrenalin receptor subtype coding sequence.
 KW Beta-2 adrenalin subtype; cyanopindrol; agonist; antagonist;
 KW asthmatic disease; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS
 FT 101..1348
 FT /*tag= a
 FT /transl_except= (pos: 338..340, aa: Xaa)
 FT /note= "Xaa-Leu, Ile"

W09735963-A1.
 PN 02-OCT-1997.
 PD 24-MAR-1997; J00982.
 PR 27-MAR-1996; JP-072914.
 PA (DAIN) DAINIPPON PHARM CO LTD.
 PI Fujii K, Furutani Y, Kawashima H, Nomura A, Yano K;
 DR WPI: 97-489627/45.
 DR P-PSDB: W34319.
 PT Novel beta-2 adrenalin receptor sub-type - useful for screening for
 PT agonists and antagonists and researching asthmatic diseases
 PS Claim 5; Page 23-26; 47pp; Japanese.
 CC This sequence encodes the protein of the invention. The protein of the
 CC invention is a beta-2 adrenalin receptor subtype with kd value of
 CC approximately 75 pM against 125I-cyanopindrol. The protein can be used in
 CC screening for agonists and antagonists, which are useful in researching
 CC asthmatic diseases.
 SQ Sequence 1400 BP; 304 A; 402 C; 363 G; 328 T;

Query Match 100.0%; Score 19; DB 1; Length 1400;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acatgacgatgccatgcc 19
 Db 221 ACATGACGATGCCATGCC 203

RESULT 5
 T93250/c
 ID T93250 standard; cDNA to mRNA; 1999 BP.
 AC T93250;
 DT 20-APR-1998 (first entry)
 DE Beta-2 adrenalin receptor subtype coding sequence.
 KW Beta-2 adrenalin subtype; cyanopindrol; agonist; antagonist;
 KW asthmatic disease; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS
 FT 190..1431
 FT /*tag= a

W09735963-A1.
 PN 02-OCT-1997.
 PD 24-MAR-1997; J00982.
 PR 27-MAR-1996; JP-072914.
 PA (DAIN) DAINIPPON PHARM CO LTD.

PI Fujii K, Furutani Y, Kawashima H, Nomura A, Yano K;
 DR MPI: 97-489627/45.
 DR P-PSDB: W34320.
 PT Novel beta-2 adrenalin receptor sub-type - useful for screening for
 PT agonists and antagonists and researching asthmatic diseases
 PS Disclosure: Page 27-30; 47pp; Japanese.
 CC This sequence encodes the protein of the invention. The protein of the
 CC invention is a beta-2 adrenalin receptor subtype with Kd value of
 CC approximately 75 pM against 125I-cyanopindolol. The protein can be used in
 CC screening for agonists and antagonists, which are useful in researching
 CC asthmatic diseases.
 SQ Sequence 1999 BP; 477 A; 513 C; 485 G; 524 T;

Query Match 100.0%; Score 19; DB 1; Length 1999;
 Best Local Similarity 100.0%; Pred. No. 0.66;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acatgacgatgccatgcc 19
 ||||||||||||||||
 DB 310 ACATGACGATGCCATGCC 292

RESULT 6
 V30468/c
 ID V30468 standard; cDNA to mRNA; 2679 BP.
 AC V30468;
 DT 14-OCT-1998 (first entry)
 DE Canine beta-2 adrenergic receptor coding sequence.
 KM Canine; beta-adrenergic receptor; brown adipose tissue; probe: human;
 KM hybridisation; ligand: ss.
 OS Canis familiaris.
 FH Key
 FT Location/Qualifiers
 FT CDS
 FT 169..1416
 FT /tag- a
 FT /product- "beta-2 adrenergic receptor"

MO975973-A2.
 PD 02-OCT-1997.
 PF 26-MAR-1997; F00537.
 PR 26-MAR-1996; FR-003730.
 PA (VERT-) VERTGEN.
 PI Drumare MF, Lenzen G, Pletzl-Rouxel F, Strosberg AD;
 DR MPI: 98-032136/03.
 DR P-PSDB: W44932.

PT Canine beta 2 and beta 3 adrenergic receptors and coding sequences -
 PT useful for identifying specific ligands and (ant)agonists to develop
 PT specific treatments for obesity in dogs
 PS Claim 1: Page 45-46; 79pp; French.
 CC This sequence represent the coding region of the canine beta 2-adrenergic
 CC receptor (RA-Ca-b2) gene. The sequence was isolated from a cDNA library
 CC constructed from polyA+ RNA purified from dog brown adipose tissue cells.
 CC The probe was a 600 bp fragment of the coding region of the human beta-3
 CC adrenergic receptor covering the region from the initiation codon to
 CC transmembrane domain 5 (TM5). The full length insert was cloned into M13
 CC for sequencing using primers V30491-V30510. The sequence can then be
 CC expressed e.g. in a mammalian cell, by subcloning into an expression
 CC vector such as pCDNA3. The beta-2 receptor can be used in comparative
 CC structure-function studies, e.g. for differential screening of ligands
 CC specific for RA-Ca-b2 or RA-Ca-b3 (W44933).
 SQ Sequence 2679 BP; 577 A; 736 C; 724 G; 642 T;

Query Match 100.0%; Score 19; DB 1; Length 2679;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acatgacgatgccatgcc 19
 ||||||||||||||||
 DB 289 ACATGACGATGCCATGCC 271

RESULT 7
 V52614/c

ID V52614 standard; cDNA; 3451 BP.
 AC V52614;
 DT 21-DEC-1998 (first entry)
 DE Human beta-2-adrenergic receptor cDNA.
 KM Human beta-2-adrenergic receptor; human; asthma; beta-agonist;
 KM polymorphism: ds.
 OS Homo sapiens.
 FH Key
 FT Location/Qualifiers
 FT CDS
 FT 1588..2829
 FT /tag- a
 FT /tag- b
 FT variation 1633
 FT /note- "A to G substitution, results in Arg16
 FT to Gly amino acid change"

MO9839477-A2.
 PD 11-SEP-1998.
 PF 26-FEB-1998; U03908.
 PR 03-MAR-1997; US-811441.
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
 PI Boushey H, Chinchilli VM, Drazen JM, Fish JE, Ford JG,
 PI Martin RJ;
 DR MPI: 98-506372/43.
 DR P-PSDB: W75777.
 PT Diagnosing asthma patients predisposed to adverse beta-agonist
 PT reactions upon regular administration - by identifying patients
 PT homozygous for allele encoding Arg at position 16 of
 PT beta2-adrenergic receptor protein
 PS Disclosure: Page 33-35; 46pp; English.
 CC This cDNA sequence codes for human beta-2-adrenergic receptor (see
 CC W75777) having an arginine residue at position 16. A novel method
 CC for identifying individuals susceptible to adverse responses to
 CC regular administration of beta-agonists comprises: (a) identifying
 CC in a genomic nucleic acid sample from the individual first and
 CC second alleles of the beta 2-adrenergic receptor gene, and (b)
 CC classifying an individual as susceptible if first and second
 CC alleles both encode Arg at residue 16 of the beta 2-adrenergic
 CC receptor protein. Beta 2-adrenergic receptor gene alleles may be
 CC identified by any known method e.g. denaturing gel electrophoresis
 CC or PCR amplification (see also V52615-17). Identification
 CC preferably comprises amplifying a portion of each allele which
 CC includes the sequence encoding residue 16, and optionally also
 CC comprises determining nucleotide sequences of these portions (e.g.
 CC by automated sequence analysis). The invention identifies a known
 CC polymorphism in the beta 2-adrenergic receptor gene as being linked
 CC to adverse responses to regular beta-agonist administration;
 CC position 16 of the encoded protein can be either Arg or Gly, and
 CC individuals homozygous for Arg16 are more susceptible.
 SQ Sequence 3451 BP; 790 A; 873 C; 895 G; 893 T;

Query Match 100.0%; Score 19; DB 1; Length 3451;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 acatgacgatgccatgcc 19
 ||||||||||||||||
 DB 1708 ACATGACGATGCCATGCC 1690

RESULT 8
 V87532
 ID V87532 standard; cDNA; 432 BP.
 AC V87532;
 DT 12-FEB-1999 (first entry)
 DE EST clone D0675.
 KM Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KM tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KM chemotaxis; chemokines; hemostasis; gene therapy; thrombolysis;
 KM receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 PD 15-OCT-1998.
 PD 15-APR-1998; U06956.

PR 10-APR-1997; US-837312.
 PA (GMY) GENETICS INST INC.
 PI Apostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D.
 PI Racie LA, Spaulding V, Treacy M;
 DR WPI: 99-070078/06.
 PT New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries
 PS Claim 1; Page 95; 641pp; English.
 CC The present sequence represents an expressed sequence tag (EST), and is
 CC a polynucleotide of the invention. The polynucleotides of the invention
 CC are all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC hematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene
 CC therapy.
 SQ Sequence 432 BP; 144 A; 73 C; 118 G; 96 T;

Query Match 91.6%; Score 17.4; DB 1; Length 432;
 Best Local Similarity 94.7%; Pred. No. 3.4;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 acatgacgatgccatgcc 19
 |||||||
 DB 113 ACATGACGATGCCATGCC 131

RESULT 9
 T97609/c
 ID T97609 standard; DNA; 2795 BP.
 AC T97609;
 DT 27-MAR-1998 (first entry)
 DE Human E2A-binding protein genomic DNA (partial sequence).
 KM E2A-binding protein; E2A-BP; human; vascular smooth muscle cell;
 KW wound healing; angiogenesis; arteriosclerosis; gene therapy; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 161..2467
 FT CDS /*tag= a
 PN WO9733900-A1.
 PD 18-SEP-1997.
 PF 14-MAR-1997; U04117.
 PR 15-MAR-1996; US-013439.
 PA (HARD) HARVARD COLLEGE.
 PI Endege WO, Haber E, Layne MD, Lee M;
 PI WPI: 97-470808/43.
 DR P-PSDB; W36815.
 PT DNA encoding human and mouse E2A binding proteins - useful for
 PT inhibiting or stimulating growth of vascular smooth muscle cells,
 PT e.g. for wound healing or treatment of rheumatoid arthritis or
 PT retinopathic diabetes
 PS Claim 14; Page 47-51; 90pp; English.
 CC This sequence represents a partial human nucleotide sequence that
 CC codes for E2A binding protein (E2A-BP, see W36815). It was
 CC isolated from a human aortic expression library using the BHLH
 CC domain of E47 as probe. E2A-BP is expressed in vascular smooth
 CC muscle cells; mRNA is preferentially expressed in the aorta.
 CC E2A-BP interacts with E2A proteins in vivo, binds to both E12
 CC and E47 and inhibits binding of E47 homodimer to an E-box probe
 CC (see T97606). A full-length DNA sequence is given in T97610, and
 CC mouse (see T97611) and rat (see T97612-13) E2A-BP sequences have
 CC also been isolated. E2A-BP nucleic acids may be used in gene
 CC therapy and antisense methods for treating vascular diseases such
 CC as arteriosclerosis, to produce transgenic or knock-out animals,
 CC as well as in methods of producing E2A-BP polypeptides that can

CC be used therapeutically to promote vascular smooth muscle cell
 CC growth e.g. for wound healing, or to identify modulator compounds
 CC suitable for treatment of e.g. arteriosclerosis and angiogenesis.
 SQ Sequence 2795 BP; 653 A; 828 C; 842 G; 472 T;

Query Match 84.2%; Score 16; DB 1; Length 2795;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 tgacgatgccatgcc 19
 |||||||
 DB 2626 TGACGATGCCATGCC 2611

RESULT 10
 T97610/c
 ID T97610 standard; cDNA; 3854 BP.
 AC T97610;
 DT 27-MAR-1998 (first entry)
 DE Human E2A-binding protein cDNA.
 KM E2A-binding protein; E2A-BP; human; vascular smooth muscle cell;
 KW wound healing; angiogenesis; arteriosclerosis; gene therapy; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1027..3564
 FT CDS /*tag= a
 PN WO9733900-A1.
 PD 18-SEP-1997.
 PF 14-MAR-1997; U04117.
 PR 15-MAR-1996; US-013439.
 PA (HARD) HARVARD COLLEGE.
 PI Endege WO, Haber E, Layne MD, Lee M;
 PI WPI: 97-470808/43.
 DR P-PSDB; W36816.
 PT DNA encoding human and mouse E2A binding proteins - useful for
 PT inhibiting or stimulating growth of vascular smooth muscle cells,
 PT e.g. for wound healing or treatment of rheumatoid arthritis or
 PT retinopathic diabetes
 PS Claim 9; Page 55-60; 90pp; English.
 CC This sequence represents a full-length nucleotide sequence that
 CC codes for human E2A binding protein (E2A-BP, see W36816). It
 CC was isolated from a human aortic RNA following 3 rounds of 5'RACE
 CC procedures. E2A-BP is expressed in vascular smooth muscle cells;
 CC mRNA is preferentially expressed in the aorta. E2A-BP interacts
 CC with E2A proteins in vivo, binds to both E12 and E47 and inhibits
 CC binding of E47 homodimer to an E-box probe (see T97606). Mouse
 CC (see T97611) and rat (see T97612-13) E2A-BP sequences have also
 CC been isolated. E2A-BP nucleic acids may be used in gene therapy
 CC and antisense methods for treating vascular diseases such as
 CC arteriosclerosis, to produce transgenic or knock-out animals, as
 CC well as in methods of producing E2A-BP polypeptides that can be
 CC used therapeutically to promote vascular smooth muscle cell growth
 CC e.g. for wound healing, or to screen for modulator compounds useful
 CC e.g. for the treatment of arteriosclerosis and angiogenesis. The
 CC E2A-BP gene promoter can be used in gene therapy methods to direct
 CC vascular smooth muscle cell-specific expression of the E2A-BP gene,
 CC antisense sequences or heterologous genes.
 SQ Sequence 3854 BP; 899 A; 1214 C; 1161 G; 580 T;

Query Match 84.2%; Score 16; DB 1; Length 3854;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 tgacgatgccatgcc 19
 |||||||
 DB 2626 TGACGATGCCATGCC 2611

RESULT 11
 O43662
 ID O43662 standard; DNA; 4131 BP.

AC 043662; (first entry)
 DE 13-OCT-1993
 KW Acetobacter cdg3 operon.
 KW Cyclic diguanylate; diguanylate phosphodiesterase;
 KW diguanylate cyclase; cellulose production; ss.
 OS Acetobacter xylinum.
 FH Key Location/Qualifiers
 FT cds 387..2606
 FT /*tag= b
 FT /label= pdeA3
 FT /note= "a diguanylate phosphodiesterase gene;
 FT the amino acid sequence R38155 cannot be
 FT deduced directly from this open reading
 FT frame due to a number of apparent
 FT (typographical ?) deletions in the printed
 FT sequence"
 FT FT 741..742
 FT /*tag= b
 FT /codon= seq: GT: aa: Val
 FT /note= "insertion of N between nucleotides 742 and
 FT 743 will restore the reading frame"
 FT FT 2522..2523
 FT /*tag= c
 FT /codon= seq: GG: aa: Trp
 FT /note= "insertion of T between nucleotides 2521 and
 FT 2522 will restore the reading frame"
 FT FT 2560..2561
 FT /*tag= d
 FT /codon= seq: CC: aa: Ala
 FT /note= "insertion of G between nucleotides 2559 and
 FT 2560 will restore the reading frame"
 FT FT 2657..4131
 FT /*tag= e
 FT /label= dgc3
 FT /note= "a diguanylate cyclase gene;
 FT the ORF is described as continuing to
 FT nucleotide 4143 although the sequence is
 FT only shown up to nucleotide 4231. The amino
 FT acid sequence R38156 cannot be deduced
 FT directly from this open reading frame which
 FT is given as printed in the specification"
 FT FT 3848..3850
 FT /*tag= f
 FT /codon= seq: GCG: aa: Ala
 FT /note= "insertion of G between nucleotides 3847 and
 FT 3848 will restore the reading frame"
 FT FT 3962..3963
 FT /*tag= i
 FT /codon= seq: CC: aa: Thr-Phe
 FT FT 109311244-A.
 FT 10-JUN-1993.
 FT 14-OCT-1992; U08756.
 FT 29-NOV-1991; US-800218.
 FT (WEYE) WEYERHAEUSER CO.
 FT PI Ben-Bassat A, Benzliman M, Calhoun RD, Gelfand DH;
 FT PI Tal R, Wong HC;
 FT WPI: 93-197062/24.
 FT DR P-PSDB: R38155, R38156.
 FT Polynucleotide sequence from Acetobacter cdg operon - encodes
 FT cyclic di:guanosine mono:phosphate degradation enzymes e.g.
 FT PS claim 3: Page 70-72; 98pp; English.
 FT CC A gene bank of Acetobacter 1306-11 was constructed in the broad host
 FT CC range, mobilisable cosmid pKT230C055. The average insert size was ca.
 FT CC 30kb. The bank was screened with a 53bp PCR amplified probe whose
 FT CC sequence was expected to be an exact match for a diguanylate
 FT CC cyclase encoding gene. Three unique cosmid clones were identified
 FT CC corresponding to three different cdg operons. The cdg3 operon
 FT CC comprises at least 2 genes. Manipulating the expression levels of
 FT CC the different cdg3 genes is contemplated, particularly in cellulose
 FT CC producing bacteria.
 FT SQ Sequence 4131 BP; 864 A; 1275 C; 1197 G; 795 T;

Query Match

81.1%; Score 15.4; DB 1; Length 4131;

Best Local Similarity 94.1%; Pred. No. 47;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 2 catgacgatgcccatagc 18
 Db 1579 CATGACGATGCCCATGC 1595
 RESULT 12
 ID 031652 standard; cDNA: 453 BP.
 AC 031652;
 DE 07-APR-1993 (first entry)
 DE Pathogenesis-related protein p14e gene isolated from tomato.
 KW PR-protein; fungicide; rust fungus; powdery mildew; ss.
 OS Lycopersicon esculentum.
 FH Key Location/Qualifiers
 FT signal_peptide 1..72
 FT mat_peptide 73..450
 FT /*tag= a
 FT /*tag= b
 FT /product= p14e
 FT FT WO9220800-A.
 FT 26-NOV-1992.
 FT 14-MAY-1992; E01063.
 FT 15-MAY-1991; GB-010544.
 FT (SANO) SANDOZ PATENT GMBH.
 FT PI Cohen Y, Guegler K, Moesinger E, Niderman T;
 FT WPI: 92-41578/50.
 FT DR P-PSDB: R28353.
 FT New P14 proteins isolated from infected tomato plants - 1e used
 FT for treating fungal infection of e.g. Uromyces appendiculatus,
 FT PT Erysiphe graminis, etc. in plants
 FT PS claim 3: Page 25; 36pp; English.
 FT CC A clone encoding PR-protein p14e (see 031652) was isolated from a
 FT CC cDNA library established from infected tomato plants. The
 FT CC library was screened for p14 genes using a probe (031657) based on
 FT CC the known p14 sequence (EMBO Journal, vol.4, #11, 2745-2749, 1985).
 FT CC The p14 clone was then used as a probe to screen a genomic library
 FT CC for other members of the p14 family. An overall total of 6 genes
 FT CC were identified, of which 5 are located in a 50kb stretch in the
 FT CC tomato genome, one of the genes appearing to be a pseudogene. The
 FT CC p14e protein is very similar to the p14e protein though has a
 FT CC stretch of 9 amino acids (residues 127 to 135) deleted from its
 FT CC C-terminal and the last 2 C-terminal amino acids (residues 125 and
 FT CC 126) substituted with respect to p14e. This gives a homology of 90%
 FT CC between the p14e and p14e mature proteins.
 FT CC See 031652-031660.
 FT SQ Sequence 453 BP; 110 A; 102 C; 119 G; 122 T;

Query Match 77.9%; Score 14.8; DB 1; Length 453;
 Best Local Similarity 88.9%; Pred. No. 69;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 catgacgatgcccatagc 19
 Db 103 CATGACGATGCCCATGC 120
 RESULT 13
 ID 031652 standard; cDNA: 480 BP.
 AC 031652;
 DE 07-APR-1993 (first entry)
 DE Pathogenesis-related protein p14e gene isolated from tomato.
 KW PR-protein; fungicide; rust fungus; powdery mildew; ss.
 OS Lycopersicon esculentum.
 FH Key Location/Qualifiers
 FT signal_peptide 1..72
 FT mat_peptide 73..477
 FT /*tag= a

```

FT      /*tag= b
FN      /product= p14a
PD      WO9220800-A.
PE      26-NOV-1992.
PF      14-MAY-1991; GB-010544.
PR      15-MAY-1991; GB-010544.
PA      (SANO ) SANDOZ PATENT GMBH.
PI      Cohen Y, Guegler K, Moesinger E, Niderman T;
PT      WPI: 92-415778/50.
DR      P-PSDB: R28350.
PT      New P14 proteins isolated from infected tomato plants - is used
PT      for treating fungal infection of e.g. Uromyces appendiculatus,
PS      Erysiphe graminis, etc. in plants
PS      Claim 3: Page 22: 36pp; English.
CC      A CDNA library was established from infected tomato plants. The
CC      library was screened for p14 genes using a probe (Q31657) based on
CC      the known p14 sequence (EMBO Journal, vol.4, #11, 2745-2749, 1985).
CC      A positive clone containing an ORF of 480bp was isolated and
CC      sequenced. It was found to encode p14a, a 159 amino acid protein
CC      differing from the previously published p14 protein by 5 additional
CC      amino acids (at AA positions 99-103, inclusive). The p14a CDNA was
CC      used to screen a genomic library for other members of the p14
CC      family. P14a has fungicidal activity and can be used to combat rust
CC      fungi and powdery mildew in various plants.
CC      See Q31652-Q31660.
SQ      Sequence 480 BP; 116 A; 110 C; 126 G; 128 T;

Query Match 77.9%; Score 14.8; DB 1; Length 480;
Best Local Similarity 88.9%; Pred. No. 70;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      2 catgacgatgccatgcc 19
DB      103 CATACGATGCCCGTCC 120
||| ||||| ||||| |||||

```

```

RESULT 14
O66939 standard; DNA; 730 BP.
AC      066939:
DT      28-JUL-1994 (first entry)
DE      GGF gene coding segment L.
KW      Glial growth factor; GGF; heregulin; mitogenesis;
KW      Schwann cell; tumour; central nervous system; ss.
OS      Homo sapiens.
FH      Key Location/Qualifiers
FT      cds 2..655 /*tag= a
FT      WO9400140-A.
FN      06-JAN-1994.
PD      29-JUN-1993; U06228.
PE      30-JUN-1992; US-907138.
PR      03-SEP-1992; US-940389.
PR      23-OCT-1992; US-965173.
PR      24-MAR-1993; US-036555.
PA      (CAMP-) CAMBRIDGE NEUROSCIENCE.
PA      (LUDW-) LUDWIG INST CANCER RES.
PI      Chen MS, Goodearl A, Hiles I, Marchionni M, Minghetti L,
PI      Stroobant P, Waterfield M;
PT      WPI: 94-025682/03.
DR      P-PSDB: R55774.
PT      Glial mitogenic polypeptide factors - useful for stimulating
PT      glial cell mitogenesis and treating glial cell tumours
PS      Claim 1-21; Fig 31; 178pp; English.
CC      wherein DNA sequence encoding a polypeptide of the formula
CC      MYBAZCX is composed of the polypeptide segments shown in
CC      sequences R55639-53 and R55766-75,
CC      wherein W comprises polypeptide segment F, or is absent;
CC      wherein Y comprises polypeptide segment E, or is absent;
CC      wherein Z comprises polypeptide segment G, or is absent; and

```

```

CC      wherein X comprises polypeptide segments C/D HKL, C/D H, C/D HL, C/D D,
CC      C/D' HL, C/D' HKL, C/D' H, C/D' D, C/D' C/D' HKL, C/D' D' HKL,
CC      C/D' D' H, C/D' D' HKL, C/D' C/D' D' H, C/D' C/D' D' HL,
CC      C/D' C/D' D' HKL, or C/D' D' HL;
CC      provided that, either
CC      (a) at least one of F, Y, B, A, Z, C, or X is of bovine origin; or
CC      (b) Y comprises polypeptide segment E; or
CC      (c) X comprises polypeptide segments C/D HKL, C/D D, HL,
CC      C/D' HKL, C/D' C/D' HKL, C/D' D, C/D' D' H, C/D' D' HL,
CC      C/D' D' HKL, C/D' D' H, C/D' D' HKL, C/D' C/D' D' H, C/D'
CC      C/D' D' HL, is claimed.
CC      DNA sequences comprising coding segments FBA, FBA', FEBA and FEBA' are
CC      also claimed
CC      The polypeptides are useful for the stimulation of glial cell
CC      (partic., Schwann cell) mitogenesis and treatment of glial cell
CC      tumours.
SQ      Sequence 730 BP; 216 A; 214 C; 159 G; 141 T;

Query Match 77.9%; Score 14.8; DB 1; Length 730;
Best Local Similarity 88.9%; Pred. No. 74;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      2 catgacgatgccatgcc 19
DB      106 CATGACGTCGTCCATGCC 123
||||| || ||||| |||||

```

```

RESULT 15
O67004 standard; DNA; 730 BP.
AC      067004:
DT      28-JUL-1994 (first entry)
DE      GGF gene coding segment L.
KW      Glial growth factor; GGF; heregulin; proliferation; antibody;
KW      Schwann cell; tumour; central nervous system; inhibition;
KW      receptor; neural; ss.
OS      Homo sapiens.
FH      Key Location/Qualifiers
FT      cds 2..655 /*tag= a
FT      WO9403644-A.
FN      17-FEB-1994.
PD      10-AUG-1993; U07491.
PE      10-AUG-1992; US-927337.
PR      25-SEP-1992; US-951747.
PR      01-DEC-1992; US-984085.
PR      29-JAN-1993; US-011396.
PA      (CAMP-) CAMBRIDGE NEUROSCIENCE INC.
PA      Gwynne DI, Marchionni M, McBurney RN;
PI      WPI: 94-065731/08.
DR      P-PSDB: R55815.
PT      Glial growth factor DNA encoding numerous polypeptide factors
PT      used for inhibiting cell proliferation - for treating carcinoma
PS      Claim 6-7, 21-24; Fig 31; 178pp; English.
CC      and nervous disorders
CC      Polypeptides of the formula
CC      VYBAZWX
CC      wherein VYBAZWX is composed of the polypeptide segments shown in
CC      sequences R4698-913 and R5807-815
CC      wherein V comprises polypeptide segment F, or is absent;
CC      wherein Y comprises polypeptide segment E, or is absent; and
CC      wherein Z comprises polypeptide segment G, or is absent; and
CC      wherein X comprises polypeptide segments C/D HKL, C/D H, C/D HL,
CC      C/D D, C/D' HL, C/D' HKL, C/D' H, C/D' D, C/D' C/D' HKL, C/D' D' HKL,
CC      C/D' D' H, C/D' D' HKL, C/D' C/D' D' H, C/D' C/D' D' HL,
CC      C/D' C/D' D' HKL, or C/D' D' HL; and
CC      C/D' D' H, C/D' D' HKL, C/D' C/D' D' H, C/D' C/D' D' HL,
CC      C/D' C/D' D' HKL, H, HL, or HKL;
CC      polypeptides comprising FBA, FBA', EBA, EBA', FEBA or FEBA' polypeptide
CC      segments are used for inhibiting proliferation of a cell, for producing

```

CC an antibody specific for a polypeptide and for detecting, in a sample,
CC the presence of a molecule capable of binding to a receptor which binds
CC to a polypeptide.
CC The polypeptides are useful for the inhibition of cell (partic.,
CC Schwann cell) proliferation. The factors are useful for the
CC treatment of neural tumours.
SQ Sequence 730 BP; 216 A; 214 C; 159 G; 141 T;

Query Match 77.9%; Score 14.8; DB 1; Length 730;
Best Local Similarity 88.9%; Pred. No. 74;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 catgacgatgccatgcc 19
||||||| || |||||
DB 106 CATGACGATGCCATGCC 123

Search completed: September 12, 2000, 23:07:20
Job time: 4102 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 12, 2000, 23:04:32 : Search time 112.94 Seconds

(without alignments)
23.141 Million cell updates/sec

Title: US-09-542-718-3

Sequence: 1 acatgacgatgccatgcc 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

- 1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/lna/5C_COMB.seq:*
- 4: /cgn2_6/prodata/1/lna/5D_COMB.seq:*
- 5: /cgn2_6/prodata/1/lna/6_COMB.seq:*
- 6: /cgn2_6/prodata/1/lna/PCBUS_COMB.seq:*
- 7: /cgn2_6/prodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	19	100.0	187	6	PCT-US91-00909-21
2	19	100.0	191	6	PCT-US91-00909-20
3	19	100.0	1242	6	PCT-US91-00909-3
4	15.4	81.1	4131	2	US-08-309-512-3
5	15.4	81.1	4131	6	PCT-US92-08756A-3
6	14.8	77.9	730	1	US-08-036-5558-77
7	14.8	77.9	730	1	US-08-469-569-77
8	14.8	77.9	730	2	US-08-249-322A-77
9	14.8	77.9	730	2	US-08-469-526A-77
10	14.8	77.9	730	3	US-08-734-591A-77
11	14.8	77.9	730	3	US-08-469-660-182
12	14.8	77.9	730	6	PCT-US94-05083C-178
13	14.8	77.9	730	6	PCT-US95-06846A-77
14	14.8	77.9	1164	1	US-07-640-476-6
15	14.8	77.9	2010	1	US-07-847-743B-12
16	14.8	77.9	2010	1	US-08-456-201-12
17	14.8	77.9	2010	3	US-08-330-161-10
18	14.8	77.9	2010	3	US-08-456-241-12
19	14.8	77.9	2010	3	US-08-440-401-10
20	14.8	77.9	2010	3	US-08-419-878B-10
21	14.8	77.9	2010	6	PCT-US92-04295A-12
22	14.8	77.9	2164	1	US-08-096-277-6
23	14.8	77.9	2164	3	US-08-550-815-6
24	14.8	77.9	2164	5	US-08-703-089-6
25	14.8	77.9	2199	1	US-07-847-743B-7
26	14.8	77.9	2199	1	US-08-096-277-7

ALIGNMENTS

27	14.8	77.9	2199	1	US-08-456-201-7	Sequence 7, App1
28	14.8	77.9	2199	3	US-08-456-241-7	Sequence 7, App1
29	14.8	77.9	2199	3	US-08-550-815-7	Sequence 7, App1
30	14.8	77.9	2199	5	US-08-703-089-7	Sequence 7, App1
31	14.8	77.9	2199	6	PCT-US92-04295A-7	Sequence 7, App1
32	14.8	77.9	3060	1	US-08-098-141-1	Sequence 1, App1
33	14.4	75.8	1898	1	US-08-063-552-12	Sequence 12, App1
34	14.4	75.8	1898	6	PCT-US93-05704-12	Sequence 12, App1
35	14.4	75.8	2228	1	US-07-726-607C-1	Sequence 1, App1
36	14.4	75.8	2228	3	US-07-843-949A-1	Sequence 1, App1
37	14.4	75.8	2228	3	US-08-218-978-1	Sequence 1, App1
38	14.4	75.8	3728	1	US-08-111-939-1	Sequence 1, App1
39	14.2	74.7	538	2	US-08-840-683-5	Sequence 5, App1
40	14.2	74.7	538	4	US-08-555-722-5	Sequence 5, App1
41	14.2	74.7	1849	4	US-08-676-166A-1	Sequence 1, App1
42	14.2	74.7	1902	1	US-07-688-352C-43	Sequence 43, App1
43	14.2	74.7	1902	4	US-08-474-379C-43	Sequence 43, App1
44	14.2	74.7	1902	5	US-09-146-249A-43	Sequence 43, App1
45	14.2	74.7	1902	6	PCT-US91-02714-40	Sequence 40, App1

RESULT 1
PCT-US91-00909-21
Sequence 21, Application PC/TUS9100909

GENERAL INFORMATION:

APPLICANT: Sledziewski, Andrzej Z.

TITLE OF INVENTION: Methods of Producing Hybrid C

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Ave.

CITY: Seattle

STATE: Washington

COUNTRY: United States

Zip: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/00909

FILING DATE: 19910208

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.408PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 187 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

PCT-US91-00909-21

Query Match 100.0%; Score 19; DB 6; Length 187;

Best Local Similarity 100.0%; Pred. No. 0.76;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 acatgacgatgccatgcc 19
|||||

Db 16 ACATGACGATGCCCATGCC 34

RESULT 2

PCT-US91-00909-20/c

Sequence 20, Application PC/TUS9100909
GENERAL INFORMATION:

APPLICANT: Sledziewski, Andrzej Z.

TITLE OF INVENTION: Methods of Producing Hybrid G

TITLE OF INVENTION: Protein-Coupled Receptors

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Ave.

CITY: Seattle

STATE: Washington

COUNTRY: United States

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/00909

FILING DATE: 19910208

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.408PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

TELEX: 3723836

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 191 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

PCT-US91-00909-20

Query Match

Best Local Similarity 100.0%; Score 19; DB 6; Length 191;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acatgacgatgcccatgcc 19
|||||
Db 176 ACATGACGATGCCCATGCC 158

RESULT 3

PCT-US91-00909-3/c

Sequence 3, Application PC/TUS9100909

GENERAL INFORMATION:

APPLICANT: Sledziewski, Andrzej Z.

TITLE OF INVENTION: Methods of Producing Hybrid G

TITLE OF INVENTION: Protein-Coupled Receptors

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Ave.

CITY: Seattle

STATE: Washington

COUNTRY: United States

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/00909

FILING DATE: 19910208

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Makl, David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.408PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

TELEX: 3723836

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1242 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1242

PCT-US91-00909-3

Query Match

Best Local Similarity 100.0%; Score 19; DB 6; Length 1242;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acatgacgatgcccatgcc 19
|||||
Db 121 ACATGACGATGCCCATGCC 103

RESULT 4

US-08-309-512-3

Sequence 3, Application US/08309512

Patent No. 5739828

GENERAL INFORMATION:

APPLICANT: Tal, Ronny

APPLICANT: Ben-Ziman, Moshe

APPLICANT: Gelfand, David H.

APPLICANT: Ben-Bassat, Arie

APPLICANT: Calhoun, Roger D.

APPLICANT: Wong, Hing C.

TITLE OF INVENTION: CYCLIC DIGUANYLATE METABOLIC ENZYMES

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 2730 Sand Hill Road

CITY: Menlo Park

STATE: California

COUNTRY: U.S.A.

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/309,512

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/800,218

FILING DATE: 29-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Bortner, Scott R.

REGISTRATION NUMBER: 34,298

REFERENCE/DOCKET NUMBER: 8145-008

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-3660
TELEFAX: (415) 854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4131 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Acetobacter xylinum
US-08-309-512-3

Query Match      81.1% Score 15.4; DB 2; Length 4131;
Best Local Similarity 94.1% Pred. No. 49;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      2 catgacgatgcccatgc 18
      ||||| ||||| |||||
Db      1579 CATGACCATGCCCATGC 1595

RESULT      5
PCT-US92-08756A-3
Sequence 3, Application PC/TUS9208756A
GENERAL INFORMATION:
APPLICANT: Tal, Ronny
APPLICANT: Benziman, Moshe
APPLICANT: Gelfand, David H.
APPLICANT: Ben-Bassat, Arie
APPLICANT: Calhoun, Roger D.
APPLICANT: Wong, Hing C.
TITLE OF INVENTION: CYCLIC DICUANYLATE
TITLE OF INVENTION: METABOLIC ENZYMES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Limbach and Limbach
STREET: 2001 Ferry Building
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08756A
FILING DATE: 19921014
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/800,218
FILING DATE: 29-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bortner, Scott R.
REGISTRATION NUMBER: 34,298
REFERENCE/DOCKET NUMBER: WEIR 20050 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-433-4150
TELEFAX: 415-433-8716
TELEX: 278356
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4131 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
```

```
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Acetobacter xylinum
PCT-US92-08756A-3

Query Match      81.1% Score 15.4; DB 6; Length 4131;
Best Local Similarity 94.1% Pred. No. 49;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      2 catgacgatgcccatgc 18
      ||||| ||||| |||||
Db      1579 CATGACCATGCCCATGC 1595

RESULT      6
US-08-036-555B-77
Sequence 77, Application US/0803655B
Patent No. 5530109
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
APPLICANT: Chen, Maio Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,555B
FILING DATE: 24-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 730
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
```

US-08-036-555B-77

Query Match 77.9%; Score 14.8; DB 1; Length 730;
Best Local Similarity 88.9%; Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 catgacgatgccatgcc 19
||||| ||| ||| |||
Db 106 CATGACGCTGCATGCC 123

RESULT 7

US-08-469-569-77
; Sequence 77, Application US/08469569
; Patent No. 5606032

GENERAL INFORMATION:

APPLICANT: Goodheart, Andrew; Stroobant, Paul;

APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

APPLICANT: Chen, Mao Su; Hiles, Ian

TITLE OF INVENTION: Glial Mitogenic Factors, Their

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESS: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,569

FILING DATE: 06-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

APPLICATION DATA: 07/940,389

FILING DATE: 03-SEP-1992

APPLICATION NUMBER: 07/907,138

FILING DATE: 30-JUN-1992

APPLICATION DATA: 07/863,703

FILING DATE: 03-APRIL-1992

APPLICATION NUMBER: U.K. 91 07566.3

FILING DATE: 10-APRIL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: LUD 5250.4

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 77:

SEQUENCE CHARACTERISTICS:

LENGTH: 730

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-469-569-77

Query Match 77.9%; Score 14.8; DB 1; Length 730;

Best Local Similarity 88.9%; Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 catgacgatgccatgcc 19
||||| ||| ||| |||
Db 106 CATGACGCTGCATGCC 123

RESULT 8

US-08-249-322A-77
; Sequence 77, Application US/08249322A
; Patent No. 5716930

GENERAL INFORMATION:

APPLICANT: Goodheart, Andrew; Stroobant, Paul;

APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

APPLICANT: Chen, Mao Su; Hiles, Ian

TITLE OF INVENTION: Glial Mitogenic Factors, Their

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESS: Felte & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/249,322A

FILING DATE: 26-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

APPLICATION DATA: 07/965,173

FILING DATE: 23-OCT-1992

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

APPLICATION DATA: 07/907,138

FILING DATE: 30-JUN-1992

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APRIL-1992

APPLICATION DATA: U.K. 91 07566.3

FILING DATE: 10-APRIL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

REFERENCE/DOCKET NUMBER: LUD 250.4

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 77:

SEQUENCE CHARACTERISTICS:

LENGTH: 730

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-249-322A-77

Query Match 77.9%; Score 14.8; DB 2; Length 730;

Best Local Similarity 88.9%; Pred. No. 80;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 catgacgaltgcacatgccc 19
||||||| || |||||
Db 106 CATGACGGTGTCCATGCC 123

RESULT 9

US-08-469-526A-77
; Sequence 77, Application US/08469526A
; Patent No. 5792849
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario Su
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; NUMBER OF SEQUENCES: 187
; PREPARATION AND USE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,526A
; FILING DATE: 06 June 1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 03-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-469-526A-77

Query Match 77.9%; Score 14.8; Db 2; Length 730;
Best Local Similarity 88.9%; Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 2 catgacgaltgcacatgccc 19
||||||| || |||||

Db 106 CATGACGGTGTCCATGCC 123

RESULT 10

US-08-734-591A-77
; Sequence 77, Application US/08734591A
; Patent No. 5854220
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Hiles, Ian
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; NUMBER OF SEQUENCES: 187
; PREPARATION AND USE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: WordPerfect (Version 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,591A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/470,335
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 03-MAR-1993
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-734-591A-77

Query Match 77.9%; Score 14.8; DB 3; Length 730;
Best Local Similarity 88.9%; Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 catgacgatgccatgcc 19
||||| |
DB 106 CATGACGATGCCATGCC 123

RESULT 11
US-08-469-660-182
; Sequence 182, Application US/08469660
; Patent No. 5876973
; GENERAL INFORMATION:
; APPLICANT: Gwynne, David I.; Marchionni, Mark;
; APPLICANT: McBurney, Robert N.
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; ZIP: 0211-2804
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,660
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/011,396
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/984,085
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/951,747
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/927,337
; FILING DATE: 10-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04585/017004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 182:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-469-660-182

Query Match 77.9%; Score 14.8; DB 3; Length 730;
Best Local Similarity 88.9%; Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 catgacgatgccatgcc 19
||||| |
DB 106 CATGACGATGCCATGCC 123

RESULT 12

PCT-US94-05083C-178
; Sequence 178, Application PC/TUS9405083C
; GENERAL INFORMATION:
; APPLICANT: Robert Sklar, Mark Marchionni,
; APPLICANT: David I. Gwynne
; TITLE OF INVENTION: METHODS FOR ALTERING
; TITLE OF INVENTION: MUSCLE CONDITION
; NUMBER OF SEQUENCES: 185
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02110-2804
; MEDIUM TYPE: Diskette, 5.25 inch, 360
; MEDIUM TYPE: kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05083C
; FILING DATE: 06-MAY-94
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,204
; FILING DATE: 08-MAR-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059,022
; FILING DATE: 06-MAY-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04585/028W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US94-05083C-178

Query Match 77.9%; Score 14.8; DB 6; Length 730;
Best Local Similarity 88.9%; Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 catgacgatgccatgcc 19
||||| |
DB 106 CATGACGATGCCATGCC 123

RESULT 13
PCT-US95-06846A-77
; Sequence 77, Application PC/TUS9506846A
; GENERAL INFORMATION:
; APPLICANT: Goodheart, Andrew David; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
; APPLICANT: Chen, Miao Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 Inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06846A
FILING DATE: 25-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/249,322
FILING DATE: 26-MAY-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, Norman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5250.5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 730
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US95-06846A-77

Query Match 77.9% Score 14.8; DB 6; Length 730;
Best Local Similarity 88.9% Pred. No. 80;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Oy 2 catgacgatgccatgccc 19
||||| || |||||
Db 106 catgacggtgtccatgcc 123

RESULT 14
US-07-640-476-6
: Sequence 6, Application US/07640476
: Patent No. 5376536
: GENERAL INFORMATION:
: APPLICANT: QUAX, WILHELMUS
: APPLICANT: LUITEN, RUDOLF G.M.
: APPLICANT: SCHURHUIZEN, PAUL W.
: APPLICANT: MRABET, NADIR
: TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
: TITLE OF INVENTION: THEIR USE
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morrison & Foerster
: STREET: 545 Middlefield Road, Suite 200
: CITY: Menlo Park
: STATE: CA
: COUNTRY: USA

```

```

ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/640.476
FILING DATE: 19910110
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kate H. Murashige
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20009.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 327-7250
TELEFAX: (415) 327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptomyces murinus
STRAIN: DSM 40091
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1164
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /EC.number=5.3.1.5
OTHER INFORMATION: /product="xylose isomerase (glucose isomerase)"
OTHER INFORMATION: /evidence=EXPERIMENTAL
OTHER INFORMATION: /standard_name="D-xylose ketol isomerase"
US-07-640-476-6

Query Match 77.9% Score 14.8; DB 1; Length 1164;
Best Local Similarity 88.9%; Pred. NO. 83;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 catgacgaagcccatgccc 19
||||| |||||||
Db 249 CATGACGAGTCCCATGCC 266

RESULT 15
US-07-847-743B-12
Sequence 12 Application US/07847743B
Patent No. 3567060
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Structure, Production and Use of
TITLE OF INVENTION: Heregulin
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/07/847,743B
FILING DATE: 19920306
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705256
FILING DATE: 24-May-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/765212
FILING DATE: 25-Sep-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/790801
FILING DATE: 08-No. 5367060-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 712P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2010 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-847-743B-12

Query Match 77.9%: Score 14.8; DB 1; Length 2010;
Best Local Similarity 88.9%: Pred. No. 87;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 catgacgatgcccatgcc 19
||||| || |||||
DB 1508 CATGACGGTGTCCATGCC 1525

Search completed: September 12, 2000, 23:04:36
Job time: 3943 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 22:46:44 : Search time 1893.64 Seconds
(without alignments)
44.246 Million cell updates/sec

Title: US-09-542-718-3

Perfect score: 19
Sequence: 1 acatgacgacgacccatgac 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*
106: gb_est87:*
107: gb_est88:*
108: gb_est89:*
109: gb_est90:*
110: gb_est91:*
111: gb_est92:*
112: gb_est93:*
113: gb_est94:*
114: gb_est95:*
115: gb_est96:*
116: gb_est97:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	19	100.0	377	80	C18250	C18250 C18250 Huma
2	17.4	91.6	522	45	A1870022	A1870022 w163h10.x
3	16.4	86.3	556	113	AO681306	AO681306 HS_5488.B
4	16.4	86.3	612	72	AM424650	AM424650 707020G08
5	16.4	86.3	800	81	C82152	C82152 C82152 Leuk
6	16.4	86.3	940	81	C82248	C82248 C82248 Leuk
7	16	84.2	214	88	R35420	R35420 y965e10.r1
8	16	84.2	266	71	AW351478	AW351478 RCL-PR007
9	16	84.2	281	81	D21603	D21603 M0573A02 mo
10	16	84.2	463	47	AL047823	AL047823 DRF2P586G
11	16	84.2	557	74	AA580751	AA580751 PM3-LT003
12	16	84.2	582	28	AA595338	AA595338 nc03h02.s
13	16	84.2	649	105	AO666234	AO666234 HS_5359.B
14	15.8	83.2	335	28	AA603161	AA603161 np55e04.s
15	15.8	83.2	341	122	FR0030532	AL026701 Fugu rubr
16	15.8	83.2	368	48	AU064622	AU064622 AU064622
17	15.8	83.2	388	46	A1958009	A1958009 f089e11.Y
18	15.8	83.2	410	72	AM421540	AM421540 f192f07.Y
19	15.8	83.2	436	123	FR0034823	AL122355 Fugu rubr
20	15.8	83.2	444	39	A1444197	A1444197 f044h11.Y
21	15.8	83.2	470	104	AO560764	AO560764 HS_2079.B
22	15.8	83.2	487	94	AO124487	AO124487 HS_3145.B
23	15.8	83.2	513	95	AO167783	AO167783 HS_2203.B
24	15.8	83.2	534	20	AA020241	AA020241 m049h11.r
25	15.8	83.2	580	123	FR0034743	AL122275 Fugu rubr
26	15.8	83.2	604	64	AM134110	AM134110 f116d03.Y
27	15.8	83.2	623	71	AM343389	AM343389 f176h11.Y
28	15.8	83.2	666	23	AA246435	AA246435 f05332.5
29	15.8	83.2	672	71	AM343362	AM343362 f176f07.Y
30	15.8	83.2	949	122	CNS00303Y	AL065380 Drosophila
31	15.4	81.1	248	27	AA495310	AA495310 fa01a12.r
32	15.4	81.1	320	50	AV147821	AV147821 AV147821
33	15.4	81.1	419	80	C28764	C28764 C28764 Rice
34	15.4	81.1	456	121	B88661	B88661 R0C11-2002
35	15.4	81.1	464	41	A1544946	A1544946 f069f09.Y
36	15.4	81.1	480	48	AU066392	AU066392 AU066392
37	15.4	81.1	484	119	AZ020513	AZ020513 R0C1-23-3
38	15.4	81.1	506	39	A1444193	A1444193 f044g04.Y
39	15.4	81.1	585	117	AO942891	AO942891 Sheared D
40	15.4	81.1	601	47	AJ272869	AJ272869 AV391075
41	15.4	81.1	608	62	AV391075	AV391075 AV391075
42	15.4	81.1	644	117	AO937984	AO937984 NB6-748R
43	15.4	81.1	748	102	AO413101	AO413101 R0C1-11-1
44	15.4	81.1	799	93	AO050823	AO050823 nbx00004c
45	15	78.9	198	24	AA327286	AA327286 EST30870

ALIGNMENTS

RESULT 1
LOCUS C18250
DEFINITION C18250 Human placenta cDNA (Tfujiwara) EST
GEN:559E06 5', mRNA sequence. Homo sapiens cDNA clone

ACCESSION C18250
VERSION C18250.1
KEYWORDS GI:1579852
SOURCE EST.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 377)
Fujiwara, T., Hirano, H., Hishigaki, H., Horie, M., Kawai, A., Kuga, Y., Kyushiki, H., Nagata, M., Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H., Suzuki, M., Takachi, A., Takeda, S., Watanabe, T., Maekawa, H., Nakamura, Y. and Takahashi, E.
Otsuka cDNA project
TITLE Unpublished (1996)
JOURNAL Contact: Tsutomu Fujiwara
COMMENT Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawachi-cho, Tokushima, Tokushima, 771-01 Japan
Tel: 0886-65-2888
Fax: 0886-37-1035.

FEATURES
source Location/Qualifiers
1..377
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GEN-559E06"
/clone.lib="Human placenta.cDNA (Tfujiwara)"
/tissue_type="placenta"
BASE COUNT 91 a 113 c 101 g 72 t
ORIGIN

Query Match 100.0%; Score 19; DB 80; Length 377;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 acatgaagatgccatgcc 19
Db 235 ACATGACGATGCCATGCC 253

RESULT 2
LOCUS A1870022
DEFINITION w163h10.x1 NCI-CGAP Brn25 Homo sapiens cDNA clone IMAGE:2429635 3'
similar to SW:NP12_MOUSE P51860 NUCLEOSOME ASSEMBLY PROTEIN 1-LIKE
2: contains element MER22 repetitive element; mRNA sequence.
ACCESSION A1870022
VERSION A1870022.1
KEYWORDS GI:5543990
SOURCE EST.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 522)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL CANCER INSTITUTE / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BRGAP), Tumor Gene Index
Unpublished (1998)
On Jan 19, 1998 this sequence version replaced gi:2286321.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldi, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LMW at:
www.bio.livl.gov/bhrp/image/image.html
Insert Length: 820 Std Error: 0.00

Db 178 CATGCCATGCCATGCC 161

RESULT 5
LOCUS C82152 800 bp mRNA EST 10-JUN-1999
DEFINITION C82152 Leukocyte of Japanese flounder 1 Paralichthys olivaceus CDNA
clone WH11-3, mRNA sequence.

ACCESSION C82152
VERSION C82152.1 GI:5039388
KEYWORDS EST.
SOURCE Paralichthys olivaceus.
ORGANISM Paralichthys olivaceus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Eurypterygii; Ctenosquamata; Acanthomorpha; Euacanthomorpha; Holacanthopterygii; Acanthopterygii; Percomorpha; Pleuronectiformes; Pleuronectoidei; Bothidae; Paralichthys.

REFERENCE 1 (bases 1 to 800)
AUTHORS Nam,B., Katagiri,T., Hong,Y., Hiroo,I. and Aoki,T.
TITLE A survey of expressed genes in the leukocyte of Japanese flounder (Paralichthys olivaceus) infected with Hirame rhadbo virus
JOURNAL Unpublished (1999)
COMMENT On May 11, 1999 this sequence version replaced gi:4776626.
Contact: Takashi Aoki
Genetics and Biochemistry
Tokyo University of Fisheries
Konan, 4-5-7, Minato 108, Japan
Tel: 03-5463-0556
Fax: 03-5463-0690
Email: ad952176s4201.tokyo-u-fish.ac.jp
clone WH11-3: similar to human KIA0174 (D79996).

FEATURES
source
1. 800
/organism="Paralichthys olivaceus"
/db_xref="taxon:8255"
/clone.lib="WH11-3"
/clone.lib="Leukocyte of Japanese flounder 1"
/sex="Female"
/cell.type="Leukocyte infected with Hirame r"
BASE COUNT 224 a 221 c 194 g 161 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 81; Length 800;
Best Local Similarity 94.4%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 catgacgatgccatgcc 19
|||||
Db 326 CATGCCATGCCATGCC 343

RESULT 6
LOCUS C82248 940 bp mRNA EST 10-JUN-1999
DEFINITION C82248 Leukocyte of Japanese flounder 1 Paralichthys olivaceus CDNA
clone WD3-5, mRNA sequence.

ACCESSION C82248
VERSION C82248.1 GI:5039484
KEYWORDS EST.
SOURCE Paralichthys olivaceus.
ORGANISM Paralichthys olivaceus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Eurypterygii; Ctenosquamata; Acanthomorpha; Euacanthomorpha; Holacanthopterygii; Acanthopterygii; Percomorpha; Pleuronectiformes; Pleuronectoidei; Bothidae; Paralichthys.

REFERENCE 1 (bases 1 to 940)
AUTHORS Nam,B., Katagiri,T., Hong,Y., Hiroo,I. and Aoki,T.
TITLE A survey of expressed genes in the leukocyte of Japanese flounder (Paralichthys olivaceus) infected with Hirame rhadbo virus

JOURNAL Unpublished (1999)
CONTACT: Takashi Aoki
Genetics and Biochemistry
Tokyo University of Fisheries
Konan, 4-5-7, Minato 108, Japan
Tel: 03-5463-0556
Fax: 03-5463-0690
Email: ad952176s4201.tokyo-u-fish.ac.jp
clone WD3-5: similar to human KIA0174 (D79996).

FEATURES
source
1. 940
/organism="Paralichthys olivaceus"
/db_xref="taxon:8255"
/clone.lib="WD3-5"
/clone.lib="Leukocyte of Japanese flounder 1"
/sex="Female"
/cell.type="Leukocyte infected with Hirame r"
BASE COUNT 255 a 258 c 211 g 216 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 81; Length 940;
Best Local Similarity 94.4%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 catgacgatgccatgcc 19
|||||
Db 311 CATGCCATGCCATGCC 328

RESULT 7
LOCUS R35420/c 214 bp mRNA EST 02-MAY-1995
DEFINITION Y965610.r1 Soares infant brain INIB Homo sapiens CDNA clone
IMAGE:38075 5', mRNA sequence.

ACCESSION R35420
VERSION R35420.1 GI:792321
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 214)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Matra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1133491.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1655
High quality sequence stops: 153 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1655 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 153.
Location/Qualifiers
1. 214
/organism="Homo sapiens"
/db_xref="GDB:410616"
/db_xref="taxon:9606"
/clone.lib="IMAGE:38075"
/clone.lib="Soares infant brain INIB"
/sex="Female"
/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"
 /note="Organ: whole brain; Vector: Lefmid BA; Site.1: Not
 I; Site.2: Hind III; 1st strand cDNA was primed with a Not
 I - oligo(dT) primer 15'
 ACTGGAAGATTCGGCCGCGAGAAATTTTCTTTTCTTTT 3';
 double-stranded cDNA was ligated to Hind III adaptors
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the Lefmid BA vector.
 Library went through one round of normalization. Library
 constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 45 a 63 c 67 g 35 t 4 others

ORIGIN

Query Match 84.2%; Score 16; DB 88; Length 214;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 tgaacatgccatgcc 19
 |||||

Db 159 TGACGATGCCATGCC 144

RESULT 8
 LOCUS AM351478 266 bp mRNA EST 01-FEB-2000
 DEFINITION RCI-BT0077-150959-001-H05 BT0077 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM351478
 VERSION AM351478.1 GI:6849191
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 266)
 HCCP http://www.ludwig.org.br/ORESTES.
 The FAPESP/LICR Human Cancer Genome Project
 unpublished (1999)
 On Jul 7, 1999 this sequence version replaced gi:5407415.
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-BT0077-
 150959-001-H05&t3=1999-09-15&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 83
 High quality sequence stop: 266.
 Location/Qualifiers

FEATURES
 source
 1..266
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT0077"
 /dev_stage="Adult"
 /note="Organ: breast; Vector: puc18; Site.1: SmaI; Site.2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

BASE COUNT 36 a 80 c 93 g 57 t

ORIGIN

Query Match 84.2%; Score 16; DB 71; Length 266;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 4 tgaacatgccatgcc 19
 |||||

Db 12 TGACGATGCCATGCC 27

RESULT 9
 LOCUS D21603/c 281 bp mRNA EST 07-OCT-1996
 DEFINITION MUS73A02 mouse embryonal carcinoma cell line F9 Mus musculus cDNA
 clone 73A02, mRNA sequence.
 ACCESSION D21603
 VERSION D21603.1 GI:618733
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 281)
 Nishiguchi, S., Joh, T., Horie, K., Zou, Z., Yasunaga, T. and Shimada, K.
 A survey of genes expressed in undifferentiated mouse embryonal
 carcinoma F9 cells: characterization of low-abundance mRNAs
 J. Biochem. 116, 128-139 (1994)
 95096008
 COMMENT On Apr 7, 1998 this sequence version replaced gi:3036588.
 Contact: Kazunori Shimada
 Department of Medical Genetics, Division of Molecular Biomedicine
 Research Institute for Microbial Diseases, Osaka University
 3-1, Yamadaoka, Suita, Osaka, 565, Japan
 Tel: 06-879-8325
 Fax: 06-879-8326.
 Location/Qualifiers

FEATURES
 source
 1..281
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="73A02"
 /clone_lib="mouse embryonal carcinoma cell line F9"

BASE COUNT 64 a 74 c 57 g 69 t 17 others

ORIGIN

Query Match 84.2%; Score 16; DB 81; Length 281;
 Best Local Similarity 88.9%; Pred. No. 4.2e+02;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 catgacatgccatgcc 19
 |||||

Db 87 CATGANGANGCCATGCC 70

RESULT 10
 LOCUS AL047823/c 463 bp mRNA EST 01-MAR-2000
 DEFINITION DKEZ586G0622_r1 586 (synonym: hutel) Homo sapiens cDNA clone
 DKEZ586G0622, mRNA sequence.
 ACCESSION AL047823
 VERSION AL047823.1 GI:4728011
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 463)
 Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
 EST (Koehler, et al.)
 unpublished (1999)
 On Jan 19, 1998 this sequence version replaced gi:2285034.
 Contact: Koehler K

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Am Kioferspitz 16a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ): Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charité, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.

No SI sequence available.
This clone (DKFZp586G0622) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers

1..463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp586G0622"
/clone_id="586 (synonym: hutel)"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT 99 a 159 c 132 g 73 t
ORIGIN

Query Match 84.2%; Score 16; DB 47; Length 463;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 tgcagatgcccatgcc 19
|||||
Db 402 TGACGATGCCCATGCC 387

RESULT 11
AM580751 557 bp mRNA EST 16-MAR-2000
LOCUS PM3-LT0031-100100-003-d04 LT0031 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM580751
VERSION AM580751.1 GI:7255800
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 557)
HGCP http://www.ludwig.org.br/ORESTES.
REFERENCE The PAPESP/LICR Human Cancer Genome Project
TITLE Unpublished (1999)
JOURNAL On Jan 6, 2000 this sequence version replaced gi:6677009.
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM3<2=PM3-LT0031-
100100-003-d04<3=2000-01-10<4=1)
Seq primer: puc 18 forward
High quality sequence stop: 504.

FEATURES

source

1..557
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="LT0031"
/dev_stage="Adult"
/note="Organ: leiomyos; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under

low stringency conditions."
BASE COUNT 115 a 169 c 174 g 98 t 1 others
ORIGIN

Query Match 84.2%; Score 16; DB 74; Length 557;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 tgcagatgcccatgcc 19
|||||
Db 547 TGACGATGCCCATGCC 532

RESULT 12
AA595338 582 bp mRNA EST 08-OCT-1997
LOCUS no35h02.s1 NCI_CGAP_Pr23 Homo sapiens cDNA clone IMAGE:1102707 3'
DEFINITION similar to TR:G607132 G607132 AEBPI MRNA. ;, mRNA sequence.
ACCESSION AA595338
VERSION AA595338.1 GI:2410688
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
REFERENCE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1453 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 440.

FEATURES

source

1..582
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1102707"
/clone_id="NCI_CGAP_Pr23"
/sex="male, pooled"
/tissue_type="prostate tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: prostate; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
5' GATTCGCGACGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.2 kb."
BASE COUNT 133 a 188 c 163 g 98 t
ORIGIN

Query Match 84.2%; Score 16; DB 28; Length 582;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 tgcagatgcccatgcc 19
|||||
Db 79 TGACGATGCCCATGCC 64

RESULT 13

A066234
LOCUS A066234 649 bp DNA GSS 23-JUN-1999
DEFINITION HS_3359_B1_E01.T7A RPECI-11 Human Male BAC Library Homo sapiens
ACCESSION A066234 genomic clone Plate=935 Col=1 Row=J, genomic survey sequence.

VERSION A066234.1 GI:5174002
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 649)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,D., Young,D., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
COMMENT On Dec 15, 1999 this sequence version replaced g1:4575425.
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPECI-11. For BAC library availability, please contact Pletier de Jong (pleterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsec.washington.edu
Plate: 935 Row: J Column: 1
Seq primer: T7
Class: BAC ends
High quality sequence stop: 649.

FEATURES
source
1..649
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Plate=935 Col=1 Row=J"
/clone_lib="RPECI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

BASE COUNT 196 a 143 c 133 g 166 t 11 others
ORIGIN

Query Match 84.2%; Score 16; DB 105; Length 649;
Best Local Similarity 100.0%; Pred. NO. 4.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 atgacgagtcacatgc 18
Db 424 ATGACGATGCCCATGC 439
|||||

RESULT 14
AA603161/c 335 bp mRNA EST 08-OCT-1997
LOCUS np55604.s1 NCI-CGAP-B1.1 Homo sapiens cDNA IMAGE:1130238 3'
DEFINITION similar to gb:D21261 SM22-ALPHA HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA603161 GI:2437022
VERSION AA603161.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 335)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITILE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced g1:2045715.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1350
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrived by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrr/image/image.html

Trace considered overall poor quality
Insert Length: 1421 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES
source
1..335
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1130238"
/clone_lib="NCI-CGAP-B1.1"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
/note="Vector: p773D-Bac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is not normalized. (The normalized version of this library is NCI-CGAP-B1.2.) Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 99 a 70 c 111 g 55 t
ORIGIN

Query Match 83.2%; Score 15.8; DB 28; Length 335;
Best Local Similarity 89.5%; Pred. NO. 5.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 acatgacgagtcacatgc 19
Db 220 ACATGATGATGCCCATGCC 202
|||||

RESULT 15
FR0030332 341 bp DNA GSS 25-JUN-1998
LOCUS FR0030332
DEFINITION Fugu rubripes GSS sequence, clone 079M1Bc12, genomic survey sequence.
ACCESSION AL026701.1 GI:3264044
VERSION AL026701
KEYWORDS GSS; genome survey sequence.
SOURCE Fugu rubripes.
ORGANISM Fugu rubripes

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Fugu. 1 (bases 1 to 341)
AUTHORS Elgar,G., Clark,M., Smith,S., Weck,S., Warner,S., Umranta,Y., Williams,G. and Brenner,S.
TITILE Direct Submission
JOURNAL Submitted (09-JUN-1998) MRC Human Genome Mapping Project Resource

Centre, Hinxton, Cambridge, CB10 1SB, UK. Email:
biohelp@hmp.mrc.ac.uk
Vector: pBluescript II KS
V_type: phagemid
PRIMER: KS

DESCR: One pass dye-terminator sequencing of cosmid cloned genomic
sequence.

FEATURES
Source Location/Qualifiers
1..341
/organism="Fugu rubripes"
/db_xref="taxon:31033"
/clone_id="cosmid_079M18"
/clone="079M18bC12"

BASE COUNT 70 a 105 c 87 g 79 t
ORIGIN

Query Match 83.2%; Score 15.8; DB 122; Length 341;
Best Local Similarity 89.5%; Pred. NO. 5.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 acatgacgatgcccatgcc 19
|||||
Db 262 ACATGATGATGCCCATGAC 280

Search completed: September 12, 2000, 22:46:49
JOB time: 3791 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:02:50 ; Search time 930.45 Seconds
(without alignments)
28.772 Million cell updates/sec

Title: US-09-542-718-4

Perfect score: 15
Sequence: 1 caatagaagcaccgc 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pal:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_scs:*
14: gb_sy:*
15: gb_un:*
16: em_fun:*
17: em_hum1:*
18: em_hum2:*
19: em_in:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_sy:*
29: em_un:*
30: em_vl:*
31: gb_hcg1:*
32: gb_hcg2:*
33: gb_in1:*
34: gb_in2:*
35: em_ba1:*
36: em_ba2:*
37: em_hum3:*
38: em_hum4:*
39: gb_pr4:*
40: gb_hcg3:*
41: gb_hcg4:*
42: gb_hcg5:*
43: gb_hcg6:*

44: gb_hcg7:*
45: em_hcg1:*
46: em_hcg2:*
47: em_hcg3:*
48: em_hum5:*
49: gb_pl3:*
50: gb_pr5:*
51: gb_hcg8:*
52: gb_hcg9:*
53: gb_hcg10:*
54: gb_hcg11:*
55: gb_hcg12:*
56: gb_hcg13:*
57: gb_hcg14:*
58: gb_in3:*
59: gb_hcg15:*
60: gb_hcg16:*
61: gb_hcg17:*
62: em_hcg4:*
63: em_hcg5:*
64: em_hcg6:*
65: em_hcg7:*
66: em_hum6:*
67: gb_hcg18:*
68: gb_hcg19:*
69: gb_hcg20:*
70: gb_hcg21:*
71: gb_hcg22:*
72: gb_hcg23:*
73: gb_hcg24:*
74: gb_hcg25:*
75: gb_hcg26:*
76: gb_hcg27:*
77: gb_hcg28:*
78: gb_hcg29:*
79: gb_hcg30:*
80: gb_hcg31:*
81: gb_vil:*
82: gb_vil2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	15	100.0	1242	11 AF022954	AF022954 Homo sapi
2	15	100.0	1242	11 AF022955	AF022955 Homo sapi
3	15	100.0	1242	11 AF022956	AF022956 Homo sapi
4	15	100.0	2305	9 HSBAR	Y00106 Human gene
5	15	100.0	3451	10 HUMADBR	M15169 Human beta-
6	15	100.0	117864	31 CEY38G4	299713 Caenorhabdi
7	15	100.0	143092	33 CEY17G7B	AL023828 Caenorhab
8	15	100.0	144368	11 AC004103	AC004103 Homo sapi
9	15	100.0	164162	40 CNS01RID	AL163152 Homo sapi
10	15	100.0	197540	72 AC023555	AC023555 Homo sapi
11	15	100.0	207139	31 CEY17G7	296049 Caenorhabdi
12	15	100.0	232489	40 AL162574	AL162574 Homo sapi
13	14.6	97.3	1286	39 AF202305	AF202305 Homo sapi
14	14	93.3	865	1 BACIS2401	M23740 B.Churtingie
15	14	93.3	865	1 BACIS2402	M23741 B.Churtingie
16	14	93.3	874	5 E07891	E07891 DNA encodin
17	14	93.3	1077	82 AF172121	AF172121 Equine in
18	14	93.3	3883	12 CGU62586	U62588 Cricetus
19	14	93.3	4140	11 HSL81779	L81779 Homo sapien
20	14	93.3	4451	5 I08083	I08083 Sequence 1
21	14	93.3	4451	5 I09103	I09103 Sequence 1
22	14	93.3	4934	5 E01676	E01676 DNA sequenc
23	14	93.3	27694	7 HASMT	D31785 Hamsenule w
24	14	93.3	37082	10 HSU230B10	Z68339 Human DNA s

25	14	93.3	53132	53	AC024290		AC024290	Homo sapi
26	14	93.3	78338	39	AF129078		AF129078	Homo sapi
27	14	93.3	84038	51	AC022713		AC022713	Homo sapi
28	14	93.3	91320	70	AC018787		AC018787	Homo sapi
29	14	93.3	93963	11	AF003528		AF003528	Homo sapi
30	14	93.3	95038	72	AC011430		AC011430	Homo sapi
31	14	93.3	98844	39	HSWRN01		AF181896	Homo sapi
32	14	93.3	100685	49	AC023754		AC023754	Homo sapi
33	14	93.3	123480	42	AC012197		AC012197	Homo sapi
34	14	93.3	129547	8	AC006434		AC006434	Genomic s
35	14	93.3	132921	10	HS82711		283850	Human DNA s
36	14	93.3	133552	69	AC015631		AC015631	Homo sapi
37	14	93.3	139350	54	AC012374		AC012374	Homo sapi
38	14	93.3	139964	75	AC019064		AC019064	Homo sapi
39	14	93.3	141602	44	AC013732		AC013732	Homo sapi
40	14	93.3	148241	41	AC009857		AC009857	Homo sapi
41	14	93.3	153689	43	AC013750		AC013750	Homo sapi
42	14	93.3	156453	32	AL138894		AL138894	Homo sapi
43	14	93.3	156945	74	AC034220		AC034220	Homo sapi
44	14	93.3	161012	51	AC018509		AC018509	Homo sapi
45	14	93.3	161136	32	AL138721		AL138721	Homo sapi

ALIGNMENTS

RESULT 1
AF022954 1242 bp DNA PRI 30-OCT-1997
LOCUS Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION AF022954
ACCESSION AF022954.1 GI:2570528
VERSION
KEYWORDS
SOURCE
ORGANISM

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Reinhaus, E., Innis, M., MacIntyre, N. and Liggett, S.B.
Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL MEDLINE
AUTHORS
TITLE
2 (bases 1 to 1242)
Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
A polymorphism of the human beta 2-adrenergic receptor within the
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL MEDLINE
AUTHORS
TITLE
3 (bases 1 to 1242)
Green, S.A., Turk, J., Innis, M. and Liggett, S.B.
Amino-terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL MEDLINE
AUTHORS
TITLE
Erratum: [[published erratum appears in Biochemistry 1994 Nov
29;33(47):14368]]
4 (bases 1 to 1242)
Liggett, S.B. and Green, S.A.
Direct Submission
Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
source
1. 1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 1242
/gene="ADRB2"
1. 1242
/gene="ADRB2"
/codon_start=1
/product="beta2-adrenergic receptor"

/protein_id="AAB82149.1"
/db_xref="GI:2570529"
/translation="MGPGNGSALLPLAPNRSHADHDVQERDEWVVGIVMSLIV
LAIVGNLVITAIKAEERLTQVNFITSLACADLVGLAVYFPGAHLTKMKMTG
NFWCEFTSIDVLCVTASIEFLCYIADRYFAITSPPRYOSLTKNKARVITLMWIV
SGISSEFLIOMHWYRATHOEAINGVANTCCDFTNOAYLAASSIVSEFVPLTMVF
YSRFOEAKROLOKIDKSEGRPHONLSQVODGRTGGLRSSKFLCKEHALKTIG
ILMGFTFCMLPEFTYVNVHVIQDLIRKEYITLLNMTGYVSGFNPLICRSPDERI
AFDELICRRSLKATGNGYSNGNTGSGSIHYEQEENKLLCEDLPGETDFVHGOG
TVPSDNDISQGNCSINDSL"
79
/gene="ADRB2"
/note="Gln27 to Glu polymorphism"
/replace="c"
BASE COUNT 276 a 330 c 326 g 310 t
ORIGIN

Query Match 100.0%; Score 15; DB 11; Length 1242;
Best local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 caatagaagccatgc 15
Db 42 CAATAGAACCATGC 56
|||||
|||||

RESULT 2
AF022955 1242 bp DNA PRI 30-OCT-1997
LOCUS Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION AF022955
ACCESSION AF022955.1 GI:2570530
VERSION
KEYWORDS
SOURCE
ORGANISM

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Reinhaus, E., Innis, M., MacIntyre, N. and Liggett, S.B.
Mutations in the gene encoding for the beta 2-adrenergic receptor
in normal and asthmatic subjects
Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
JOURNAL MEDLINE
AUTHORS
TITLE
2 (bases 1 to 1242)
Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
A polymorphism of the human beta 2-adrenergic receptor within the
fourth transmembrane domain alters ligand binding and functional
properties of the receptor
J. Biol. Chem. 268 (31), 23116-23121 (1993)
JOURNAL MEDLINE
AUTHORS
TITLE
3 (bases 1 to 1242)
Green, S.A., Turk, J., Innis, M. and Liggett, S.B.
Amino-terminal polymorphisms of the human beta 2-adrenergic
receptor impart distinct agonist-promoted regulatory properties
Biochemistry 33 (32), 9414-9419 (1994)
JOURNAL MEDLINE
AUTHORS
TITLE
Erratum: [[published erratum appears in Biochemistry 1994 Nov
29;33(47):14368]]
4 (bases 1 to 1242)
Liggett, S.B. and Green, S.A.
Direct Submission
Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
source
1. 1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 1242
/gene="ADRB2"
1. 1242
/gene="ADRB2"
/codon_start=1
/product="beta2-adrenergic receptor"

```

/product="beta2-adrenergic receptor"
/protein_id="AAB82150.1"
/db_xref="GI:2570531"
LAI|FBIANLVTTTAAKEERLQGTAVNYETFTSLACADLVNGLAVPFGAAHILMKMWTFC
NMEEPTNSIDVLCVTASIERFLCYIADVPAISPEKYOGLTKRKARVILLMWYIVLWVF
SGLSFPDIOHWYRATHOEAINGYANTCCDEPTNOAYLAASSISFYVLVYVWFV
YSRPFOEAKROLOKIDKSEGFHONLSOVODRGRTGRLRSSFCLEKHAKLTQD
IKNTFTPLCHLPFFIVNVHYIONLRLREYITLLNIGCVNSGNPDLTKRSDFDR
AFQBLCLURSSLSAAGNGYSSNGNTGEOSGYHVEOEKENLLCEDLPGETDEFVGHQC
TVPDNDINDSOGRCNSTDSL"
100
/gene="ADR82"
/note="val34 to Met polymorphism"
/replacement="g"
BASE COUNT      277 a      331 c      324 g      310 t
ORIGIN
Query Match          100.0%   Score 15:   DB 11:   Length 1242:
Best Local Similarity 100.0%   Pred. NO. 2,1e+02:
Matches 15:  Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Oy      1 caatagaagaccatgc 15
         |||||||||
Db       42 CAATTGAAGCCTATGC 56
RESULT 3
AF022956      LOCUS             AF022956      1242 bp      DNA              PRI           30-OCT-1997
DEFINITION     Homo sapiens beta2-adrenergic receptor (ADR82) gene, complete cds.
ACCESSION      AF022956
VERSION        AF022956.1 GI:2570532
KEYWORDS
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
               Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE       1 (bases 1 to 1242)
AUTHORS        Reihnsaus, E., Innis, M., Macintyre, N. and Liggett, S.B.
TITLE          Mutations in the gene encoding for the beta 2-adrenergic receptor
               in normal and asthmatic subjects
JOURNAL        Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE        93192047
REFERENCE       2 (bases 1 to 1242)
AUTHORS        Green, S.A., Cole, G., Jacinto, M., Innis, M. and Liggett, S.B.
TITLE          A polymorphism of the human beta 2-adrenergic receptor within the
               fourth transmembrane domain alters ligand binding and functional
               properties of the receptor
JOURNAL        J. Biol. Chem. 268 (31), 23116-23121 (1993)
MEDLINE        94043092
REFERENCE       3 (bases 1 to 1242)
AUTHORS        Green, S.A., Turki, J., Innis, M. and Liggett, S.B.
TITLE          Amino-terminal polymorphisms of the human beta 2-adrenergic
               receptor impart distinct agonist-promoted regulatory properties
JOURNAL        Biochemistry 33 (32), 9414-9419 (1994)
MEDLINE        94347707
REMARK         Erratum: [[published erratum appears in Biochemistry 1994 Nov
                29;33(47):1145681]]
                4 (bases 1 to 1242)
                Liggett, S.B. and Green, S.A.
                Direct Submision
                Submitted (04-SEP-1997) Medline, Univ of Cincinnati, 231 Bethesda
                Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
Source
gene
cds
/gene="ADR82"
1..1242
/location/Qualifiers
1..1242
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..1242
/gene="ADR82"
1..1242
/gene="ADR82"

```

```

/codon_start=1
/product="beta2-adrenergic receptor"
/protein_id="AAB82151.1"
/db_xref="GI:2570533"
/translation="MGQPGNGSAFLAPNRSHPADHDVTQORDEYVVVGMKGVLSLIV
LAIEGVNLTVALAKFERLOTVTNFTSTLACADLVGLAVVPEGAHILMKMTGEG
NMFQPMFTSIDVLCVTASIEFLCYAVADRAFAITSPKYSOLILKNKRAVILIMVIV
SGLSPLPIQHMWRATHOELINCYANETCCDFPTNOAYAIASSIVSYVPLVMVY
YSRVEQAKROLOKIDSSEGRFHQVNSQVEDRGTGHLRSSRFCLKERKAKLTJGA
IMOTFTLCWLPFPFIVNIVHIDNLIRKVEYILNMIGVNSFPNLIYCRSPDFRI
AFQELLCLRRSLKAVNGSYNSNGTGEOSGYHVEOEKENKLCEDLPGTEDFVGHQGG
TVPSDNIIDSGRNCSTNDSL"
491
/feature="ADRB2"
/feature="Thr164 to Ile polymorphism"
/replacement="C"
BASE COUNT 276 a 330 c 325 g 311 t
ORIGIN
variation
Query Match 100.0%; Score 15; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 caatagaagccatgc 15
|||||
Db 42 CAATAGAAAGCCATGC 56
|||||
RESULT 4
HBBAR 2305 bp DNA PRI 12-SEP-1993
LOCUS
DEFINITION Human gene for beta-adrenergic receptor (beta-2 subtype).
VERSION Y00106
KEYWORDS beta-adrenergic receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2305)
Schofield, P.R., Rhee, L.M. and Peralta, E.G.
Primary structure of the human beta-adrenergic receptor gene
Nucleic Acids Res. 15 (8), 3656 (1987)
87203400
2 (bases 1 to 2305)
Schofield, P.R.
Direct Submission
Submitted (20-OCT-1987) to the EMBL/GenBank/DBJ databases
1. 2305
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="Manialis human"
/clone="lambda/betaARI7"
794. 2035
/feature="beta-adrenergic receptor (AA 1 - 413)"
/codon_start=1
/protein_id="CAA68289.1"
/db_xref="GI:29371"
/db_xref="SWISS-PROT:P07550"
/translation="MGQPGNGSAFLAPNRSHPADHDVTQORDEYVVVGMKGVLSLIV
LAIEGVNLTVALAKFERLOTVTNFTSTLACADLVGLAVVPEGAHILMKMTGEG
NMFQPMFTSIDVLCVTASIEFLCYAVADRAFAITSPKYSOLILKNKRAVILIMVIV
SGLSPLPIQHMWRATHOELINCYANETCCDFPTNOAYAIASSIVSYVPLVMVY
YSRVEQAKROLOKIDSSEGRFHQVNSQVEDRGTGHLRSSRFCLKERKAKLTJGA
IMOTFTLCWLPFPFIVNIVHIDNLIRKVEYILNMIGVNSFPNLIYCRSPDFRI
AFQELLCLRRSLKAVNGSYNSNGTGEOSGYHVEOEKENKLCEDLPGTEDFVGHQGG
TVPSDNIIDSGRNCSTNDSL"
809. 817
/feature="N-linked glycosylation site"
836. 844
/feature="N-linked glycosylation site"
misc_feature
misc_feature
misc_feature

```

misc_feature 896..967
 /note="membrane spanning domain I"
 misc_feature 1007..1078
 /note="membrane spanning domain II"
 misc_feature 1114..1180
 /note="membrane spanning domain III"
 misc_feature 1247..1315
 /note="membrane spanning domain IV"
 misc_feature 1385..1450
 /note="membrane spanning domain V"
 misc_feature 1616..1687
 /note="membrane spanning domain VI"
 misc_feature 1712..1774
 /note="membrane spanning domain VII"

BASE COUNT 495 a 616 c 649 g 545 t
 ORIGIN

Query Match 100.0%; Score 15; DB 9; Length 2305;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatgc 15
 |||||||
 Db 835 CAATGAGGCGATGC 849

RESULT 5
 HUMADBR 3451 bp mRNA PRI 13-FEB-1996
 LOCUS Human beta-2-adrenergic receptor mRNA, complete cds.
 DEFINITION M15169 J02728 M16106
 ACCESSION M15169.1 GI:178201
 VERSION M15169.1
 KEYWORDS adrenergic receptor.
 SOURCE Homo sapiens (clone: pTF.) (tissue library: Evan Sadler) placenta cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 3451)
 AUTHORS Koblika,B.K., Dixon,R.A., Fritelle,T., Dohman,H.G., Bolanowski,M.A.,
 Dixon,R.A., Keller,P., Caron,M.G. and Lefkowitz,R.J.
 TITLE Delineation of the intronless nature of the genes for the human and hamster beta-2-adrenergic receptor and their putative promoter regions

JOURNAL J. Biol. Chem. 262 (15), 7321-7327 (1987)

MEDLINE 87222358

REFERENCE 2 (bases 1399 to 1985)
 AUTHORS Koblika,B.K., Dixon,R.A., Fritelle,T., Dohman,H.G., Bolanowski,M.A.,
 Sigal,I.S., Yang-Feng,T.L., Francke,U., Caron,M.G. and Lefkowitz,R.J.

TITLE cDNA for the human beta-2-adrenergic receptor: a protein with multiple membrane-spanning domains and encoded by a gene whose chromosomal location is shared with that of the receptor for platelet-derived growth factor

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84 (1), 46-50 (1987)

MEDLINE 87092393

FEATURES Location/Qualifiers

1..3451
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="pTF."
 /tissue_type="Placenta"
 /tissue_lib="Evan Sadler"
 /map="5q31-q32"
 1369..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.): G00-120-541"
 1369..3383
 /gene="ADRB2"
 1376..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.): G00-120-541"

MRNA

1379..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.): G00-120-541"
 1388..3383
 /gene="ADRB2"
 /note="b-2-adr mRNA (alt.): G00-120-541"
 1487..1546
 /gene="ADRB2"
 /note="putative"

CDS

/codon_start=1
 /db_xref="GDB:G00-120-541"
 /protein_id="AA88014.1"
 /db_xref="GI:560761"
 /translation="MRLPGVSRPAEPRGSGAR"
 1588..2829
 /gene="ADRB2"

CDS

BASE COUNT 790 a 873 c 895 g 893 t
 ORIGIN

Query Match 100.0%; Score 15; DB 10; Length 3451;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatgc 15
 |||||||
 Db 1629 CAATGAGGCGATGC 1643

RESULT 6
 CEY38G4/c
 LOCUS
 DEFINITION Caenorhabditis elegans chromosome II clone Y38G4, *** SEQUENCING IN PROGRESS ***; in unordered pieces.
 ACCESSION 299713
 VERSION 299713.1 GI:2467102
 KEYWORDS HTG; HTGS PHASE1.
 SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdittia; Rhabdittida;
 Rhabdittina; Rhabdittidae; Peloderiinae; Caenorhabditis.
 1 (bases 1 to 117864)
 Sulston,J.

TITLE Direct Submission
 JOURNAL Submitted (03-OCT-1997) Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10 1HQ, UK and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jae@esanger.ac.uk or rwenematode.wustl.edu
 On Oct 7, 1997 this sequence version replaced gi:2464982.
 IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress

and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated
 with foreign sequence from E.coli, yeast, vector, phage etc. Order of segments is not known; 800 n's separate segments.
 * NOTE: This is a 'working draft' sequence.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

* be preserved.

FEATURES
 source 1..117864
 /organism="Caenorhabditis elegans"
 /db_xref="taxon:6239"
 /chromosome="II"
 /clone="Y38G4"

BASE COUNT 26604 a 15058 c 14789 g 26050 t 35363 others
 ORIGIN

Query Match 100.0%; Score 15; DB 31; Length 117864;
 Best local similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 caatagaagcattcgc 15
 |||
 Db 108842 CAATAGACCATGC 108828

RESULT 7
 CEY17G7B 143092 bp DNA INV 26-JAN-2000
 LOCUS
 DEFINITION Caenorhabditis elegans cosmid Y17G7B, complete sequence.
 ACCESSION AL033828.296049
 VERSION AL033828.1 GI:3217816
 KEYWORDS
 SOURCE
 ORGANISM
 Caenorhabditis elegans.
 Caenorhabditis elegans.
 Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 Rhabditina; Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
 1 (bases 1 to 143092)
 none.
 Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium
 Science 282 (5396), 2012-2018 (1998)
 99069613
 JOURNAL
 MEDLINE
 REMARK
 The C. elegans Sequencing Consortium.
 Erratum: [published errata appear in Science 1999 Jan
 1;283(5398):35 and 1999 Mar 26;283(5410):2103 and 1999 Sep
 3;285(5433):14931]
 2 (bases 1 to 143092)
 Smye,R.
 DIRECT SUBMISSION
 Submitted (09-JUN-1998) Nematode Sequencing Project, Sanger Centre,
 Hinxton, Cambridge CB10 1RO, England and Department of Genetics,
 Washington University, St. Louis, MO 63110, USA. E-mail:
 jess@sanger.ac.uk or twenematode.wustl.edu
 Coding sequences below are predicted from computer analysis, using
 predictions from GeneFinder (P. Green, U. Washington), and other
 available information.
 For a graphical representation of this sequence and its analysis
 see:-
 http://webcace.sanger.ac.uk/cg1-
 bin/display?db=wormbase&class=Sequence&object=Y17G7B
 Current sequence finishing criteria for the C. elegans genome
 sequencing consortium are that all bases are either sequenced
 unambiguously on both strands, or on a single strand with both a
 dye primer and dye terminator reaction, from distinct subclones.
 Exceptions are indicated by an explicit note.
 IMPORTANT: This sequence is NOT necessarily the entire insert of
 the specified clone. It may be shorter because we only sequence
 overlapping sections once, or longer because we arrange for a small
 overlap between neighbouring subclones.
 IMPORTANT: This sequence is not the entire insert of clone Y17G7B.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring subclones.
 The true right end of clone W03C9 is at 100 in this sequence. The
 true right end of clone Y17G7 is at 143092 in this sequence. The
 start of this sequence (1..100) overlaps with the end of sequence
 26516.
 The end of this sequence (142993..143092) overlaps with the start

of sequence AL117199.

FEATURES
 source 1..143092
 /organism="Caenorhabditis elegans"
 /db_xref="taxon:6239"
 /chromosome="II"
 /clone="Y17G7B"
 complement(3569..4874)
 /gene="Y17G7B.1"
 complement(join(3569..3750,3804..3945,4851..4874))
 /gene="Y17G7B.1"
 /note="similar to Acyl CoA binding protein"
 /codon_start=1
 /protein_id="CA119448.1"
 /db_xref="SPTREMBL:O9XXJ2"
 /db_xref="GI:3947596"
 /translation="MPSLEKKMTGLTFEIAAEEMRLKSEPTRERLKLYALYKQA
 LHGDIPNEDVYVPVPAQDEYGRKKYAAWMSQKANSKCRADYVAIAEEMIKKYGNNIV
 RKMNSEWVSVDY"
 11683..19658
 /gene="Y17G7B.2c"
 11683..19658
 /gene="Y17G7B.2b"
 11683..19658
 /gene="Y17G7B.2a"
 join(11683..11753,11836..11928,11974..12304,13242..13462,
 14218..14331,15865..16127,17187..17299,18228..18428,
 18624..18714,19438..19658)
 /gene="Y17G7B.2b"
 /note="cDNA EST EMBL:J35727 comes from this gene"
 /codon_start=1
 /protein_id="CA119468.1"
 /db_xref="GI:3947616"
 /db_xref="SPTREMBL:O9XXH4"
 /translation="MRSSKGGGRQAAPKTAFTYCYCDGKRELSEVEVCSCLKMF
 HORCKEPEFERNSSNGVPEMICYPTFCOCQPTADMKAKKADLVOMCVLATLSA
 ERLKADGKSAHVPEDEFTVLSLKDEIVYNNEMNTLAIKOKEMHONLAPTLIKE
 KNIPVONHDDDLFALAEKNLSLGLPHEAVVILIGRPILERENREPRHIELPIESPK
 TNGASRRHAEAPVPGKKOKLADVSTAPAGVOVDIPFSKDNRYVLTEDVPNVE
 DPAMNOSSAEVPVIFSFHYRELNPVANSNDRAFOQLINSGNSITGFGYMARASH
 GVSKGTVFEVNEVDOPDDSHIRIGSQSVASLQCVNKSYSGRSKHGKFIHEAK
 GKHYHGFGKGDVGLCLHLPDKKLOIPANLPSEKYPVSHKGNLSIFRANFEE
 VOESADIAKTLVEMPGSYIEEFHNKSCGKAYENYACAYPSISIPSSATATNMG
 PKRNLPRCATGTHARADEQOHEDTLDMLYLVSKEVNLDPHPRVAREDDDDKDKK
 EIKQET"
 join(11683..11753,11836..11928,11980..12304,13242..13462,
 14218..14331,14622..14673,18228..18428,18624..18714,
 19438..19658)
 /gene="Y17G7B.2c"
 /note="cDNA EST EMBL:D74713 comes from this gene; cDNA EST
 YK470e5.3 comes from this gene; cDNA EST YK470e5.5 comes
 from this gene"
 /codon_start=1
 /protein_id="CA119469.1"
 /db_xref="GI:3947617"
 /db_xref="SPTREMBL:O9XXH3"
 /translation="MRSSKGGGRQAAPKTAFTYCYCDGKRELSEVEVCSCLKMF
 HORCKEPEFERNSSNGVPEMICYPTFCOCQPTADMKAKKADLVOMCVLATLSER
 LKADGKLSAEHVPEDEFTVLSLKDEIVYNNEMNTLAIKOKEMHONLAPTLKSN
 IFVONHDDDLFALAEKNLSLGLPHEAVVILIGRPILERENREPRHIELPIESPKR
 GASKRHAEAPVPGKKOKLADVSTAPAGVOVDIPFSKDNRYVLTEDVPNVEP
 IIPFLAEFLYIEFTGCVGLCLHLPDKKLOIPANLPSEKYPVSHKGNLSIFK
 ANFEEVOESADIAKTLVEMPGSYIEEFHNKSCGKAYENYACAYPSISIPSSAT
 ATNMLPKRNLPRCATGTHARADEQOHEDTLDMLYLVSKEVNLDPHPRVAREDDDD
 VDIKKEIKQET"
 join(11683..11753,11836..11928,11980..12304,13242..13462,
 14218..14331,15865..16127,17187..17299,18228..18428,
 18624..18714,19438..19658)
 /gene="Y17G7B.2c"
 /note="Y17G7B.2a"
 /codon_start=1
 /protein_id="CA119449.1"
 /db_xref="GI:3947597"

```

/translation="MRSSKGGRRGROAAPTATTCVOCODKRELSVEVYCTCLMKE
HGKLEFHEHNSNVPMICTCTCTCOCPPAEDMKAKKADJOMCIVUATLTSER
LKADGKLESEHVEDPTVLSLKEIVPMENMYMTALIKOKEHONIAPLILEKRN
IFVOHNDDEDFALAEKRLSLGPIHEAVKLGKRIEENNEPRIEELPIEGEETR
GAKRRHAEPVTKKOKLAADYSSTAANGVOIDIPKENDRYLTLEVDNAPDE
AMNONSAAVIVSPFRELINPTVANSYSLACVQKNGKFSYSGMSRGTKEHAKG
SKGTWYFEVNDOPDPSHIRIGMSOVSASLOACVGNKFSYSGMSRGTKEHAKG
KIHFGGKQGDVIGCLILPVDKLIOPALPSEKFIIPYSHKGFNLISFRANFEVQ
EESADLAKTLVEMPGSITIEFFNGKSGCRKENTYAGAYPSISIFKSATATNMLGPK
FRNLPRGATGIHARADEQOHEOTLSDMLIVSEVNLHDHPRKRRDDDDVDIKKEI
KOET"
gene      20244..23044
/gene="Y1767B.3"
join(20244..23081,21383..21553,22571..23044)
/gene="Y1767B.3"
/note="similar to Metallo-beta-lactamase superfamily"
/codon_start=1
/protein_id="CA19450.1"
/db_xref="GI:3947598"
/translation="MYYVKSLLRRADNFYTVKSSSEARALLVDLYNEDEYELADK
ENIDITAVITLHHHHDGCGNEGRROFPVPMVIGDGRIPADRHVKGHGMAFGL
OIKLSTPCHTSGHCHITRPADSTSPCVITGDFLFIAGGREGFETAPOMDVAL
NEILKNLPEVOIFPGHEIVANLKFCHVPEPKRAQKLEMAQRQIEGGFTVST
VAERKATPEMRRESEIOLSIGTCAVVGMAKLRMKKMF"
24223..29110
/gene="Y1767B.4"
join(24223..24435,25013..25162,26304..26685,28307..28435,
28669..29110)
/gene="Y1767B.4"
/note="cDNA EST EMBL:D73584 comes from this gene; cDNA EST
YK532a10.3 comes from this gene; cDNA EST YK539f8.3 comes
from this gene"
/codon_start=1
/protein_id="CA19451.1"
/db_xref="GI:3947599"
/translation="WSTVEAASQEDIALAGAVLYKSCPDGSIPIRGDFSTAG
PDFSLAISLSTMTSGQATHLAQAIOQVNMLSLRDPTCDDEKIFPYREGROKR
SCITLGYSTNLVSTGLREVLYCQVRNLVDCIYVSAGLIEDLCKLKESTLFTFM
DGAKRSGSNMNAKAVLIPNDNYCAFEDMLPILDECEVEEHLNMTPEKLRGDS
RIDGESSITLWAKHRIIPFCPLDGLSLDMLTFHVSCKSPGLRVDIVDVRHINT
TAVKSPKSGITLGGGVYKHHNNANLRNGADHYVYINTGCEFDGSDSCAQPDPAVS
WGVKRPAGAVVHAHEATLPEPLIYAEFFAHHEGKD"
complement(32881..37816)
/gene="Y1767B.5"
complement(join(32881..33253,34402..35645,36629..37493,
37653..37816))
/gene="Y1767B.5"
/note="similar to MCW2/3/5 family; cDNA EST EMBL:T02370
comes from this gene; cDNA EST YK244e4.5 comes from this
gene; cDNA EST YK407d8.3 comes from this gene"
/codon_start=1
/protein_id="CA19452.1"
/db_xref="GI:3947600"
/translation="MADRANNDVDYDROPLIADADDDVDSIDEMFNNEDEPDEE
GENLFEDDMEVYREQELDQYSESQMDASVSGISVSARRAAREMAOROLLDD
ALMYEDGDEEDVYTRRRGRGRGDAADDSPMEEDIPVILENITGRIRHVS
DEAVAKEIRERENLRFHEPKNOKTYIOMIKSMAANRSLSEFVDSLEFDDGBO
NTSYFLPEARNEMLAINDRAATEVVMNMYFMRVCNEIKVVISOLPEEDIRMLROY
HLMMLRTAGVYTIASGILPOLAVVYDCVAGCYLGPVQONDEDEVRPTIOPSCGK
GPELVENYVYHNTYRTIMQESPNKVAAGRLPRSKDVILLDLDCSKPGEIEVYG
VYTNNDGSLNTRKQFPVFTLHANHTNKKMASDOLITDDIRAIIRLSOPKASQEL
RVFSIAPSLYIGHDVYKALALALFRGEAKNPGAKRNLRLGAVNLADKVCGLIDEDK
RYAAHIAPRSIVLTGOGASAVGLTAVGRHPVRENTLRGAVNLADKVCGLIDEDK
MSDODRTSHEAEQOOSISISKAGIYLSHARCTVIAASVIGRNLPRTEAEVNDL
TEPILSRFDVLYEIRDSYSDVEDERLAKVGNHRTNHDNDAKIVGESLEDDOME
RTGVRILPODLILKXIIYAREKCHPTLPBOHSEKFSFNIAQKSMEMAGSAVITVRH
VESKIRLSEAHAKLILRSYVNDCCAAALIRVMEISYVNOAKSIMMKKTFSRILTE
NRSANELLILTLQOLIRQOMHYTARAACCTILQSVTIESESEIEKAQOLRIENVVF
YTSIEFASNNELVDPSTKTIQVEIF"
52521..54875

```

CDS

/gene="Y1767B.6"
 join(52521..52555,53285..53390,53728..53819,53916..54102,
 54595..54744,54795..54875)

Query Match

Best Local Similarity 100.0%; Score 15; DB 33; Length 143092;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caataagaacatgc 15
 Db 39325 CAATAGAACCATGC 39339

RESULT 8

AC004103/c DNA PRI 18-APR-1998
 LOCUS Homo sapiens Xp22 BAC GS-619J3 (Genome Systems Human BAC library)
 DEFINITION complete sequence.

AC004103 GI:3063487

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The repeat regions shown were identified using RepeatMasker by Adrian Smil.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

Location/Qualifiers
 1..144368
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="GS-61933"
/chromosome="X"
/clone_11b="Genome Systems Human BAC library"
/map="Xp22"
complement(1..358)
repeat_region
/rpt_family="LRR7"
complement(380..400)
repeat_region
/rpt_family="AT-rich"
complement(843..1139)
repeat_region
/rpt_family="AluSg"
1389..1507
/rpt_family="(GAA)n"
complement(2262..2536)
repeat_region
/rpt_family="LTR16c"
2905..2995
repeat_region
/rpt_family="MIR"
3019..3241
/rpt_family="MER58A"
complement(3366..3465)
repeat_region
/rpt_family="L2"
3601..3800
/rpt_family="MER69B"
complement(3871..4170)
repeat_region
/rpt_family="AluY"
complement(4800..5035)
repeat_region
/rpt_family="MIR"
complement(6283..6516)
repeat_region
/rpt_family="L1MA4"
6517..6816
repeat_region
/rpt_family="AluJo"
6817..6850
/rpt_family="AT-rich"
complement(6851..7385)
repeat_region
/rpt_family="L1MA6"
complement(8269..8773)
repeat_region
/rpt_family="L1ME1"
8878..9769
/rpt_family="L1ME3"
9785..10096
/rpt_family="AluSg"
10097..10169
/rpt_family="L1ME3"
complement(10741..10770)
repeat_region
/rpt_family="(CA)n"
complement(12086..12112)
repeat_region
/rpt_family="AT-rich"
12204..12628
/rpt_family="MSTRB"
13903..14308
/rpt_family="MER4D"
complement(14324..14807)
/Note="Region: Similar to gb:X69391.60S ribosomal protein
L6"
repeat_region
15219..15790
/rpt_family="MER4D"
16692..16784
/rpt_family="MIR"
complement(17077..17123)
repeat_region
/rpt_family="(CA)n"
complement(18203..18330)
repeat_region
/rpt_family="MIR"
18351..18884
/rpt_family="MLT1E"
/rpt_family="L1PA2"
complement(19499..19645)
repeat_region
/rpt_family="L2"
19840..20445
/rpt_family="L1MA2"
complement(21455..22132)
repeat_region
/rpt_family="L1ME3A"
/rpt_family="L1ME3"
complement(22671..22766)
/rpt_family="TIGGER2"
22771..23204

```

```

repeat_region
complement(23230..23597)
/rpt_family="TIGGER2"
23587..23655
repeat_region
/rpt_family="MER8"
25665..25970
/rpt_family="AluY"
26621..27362
STS
/standard_name="DXS43"
repeat_region
complement(29197..29351)
/rpt_family="MER34"
repeat_region
complement(30502..30607)
/rpt_family="MER5A"
complement(30613..30914)
/rpt_family="AluSx"
complement(30919..31035)
/rpt_family="MIR"
3178..31867
/rpt_family="MER34"
complement(32341..32724)
/rpt_family="MSTRB"
repeat_region
complement(32786..33167)
/rpt_family="LRR7"
repeat_region
complement(33168..33729)
/rpt_family="HERVH"
complement(33723..33983)
/rpt_family="HERVH"
complement(33974..34568)
/rpt_family="HERVH"
complement(34589..35228)
/rpt_family="HERVH"
complement(35229..37784)
/rpt_family="HERVH"
repeat_region
complement(37785..37962)
/rpt_family="LRR7"
repeat_region
complement(38041..38170)
/rpt_family="LRR7"
complement(38620..40203)
/rpt_family="MSTRB-internal"
complement(40205..40615)
/rpt_family="MSTRB"
complement(40672..41325)
/rpt_family="MER31-internal"
41371..41804
/rpt_family="L1MB8"
complement(42696..42722)
/rpt_family="AT-rich"
complement(42789..42867)
/rpt_family="MER31-internal"
43140..43337
/rpt_family="CHARLE1"
43780..44186
/rpt_family="LRR7"
complement(44227..44318)
/rpt_family="FLAM_A"
44447..44480
/rpt_family="U2"
45231..45407
/rpt_family="MIR"
complement(46263..46611)
/rpt_family="L1M4"
complement(46607..52764)
/rpt_family="L1PA2"
complement(52787..52832)
/rpt_family="MIR"
complement(53193..53543)
/rpt_family="MLT1A1"
complement(54583..54731)
/rpt_family="MER5A"

```

Query Match 100.0%; Score 15; DB 11; Length 144366;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatagaagccatgc 15
 DB 120217 CAATAGAAGCCATGC 120203

RESULT 9
 CENS01RID 164162 bp DNA HTG 04-APR-2000
 LOCUS Homo sapiens chromosome 14 clone R-98N22, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 5 ordered pieces.
 ACCESSION AL163152
 VERSION AL163152.1 GI:7452890
 KEYWORDS HTG: HTGS_PHASE2; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 (bases 1 to 164162)
 Genoscope.
 Direct Submission
 Submitted (04-APR-2000) to the EMBL/GenBank/DBJ databases
 IMPORTANT: This sequence is unfinished and does not necessarily
 represent the correct sequence. Work on the sequence is in progress
 and the release of this data is based on the understanding that the
 sequence may change as work continue. The sequence may be
 contaminated with foreign sequence from E.coli, yeast, vector,
 phage, etc.
 Contig order : 6 5 9 7 8, 1000 N's separate segments Contig 6 :
 Length 12418 bp
 Contig 5 : length 7792 bp
 Contig 9 : length 93944 bp
 Contig 7 : length 25154 bp
 Contig 8 : length 20854 bpXX.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 12418: contig of 12418 bp in length
 * 12419 13418: gap of 1000 bp
 * 13419 21210: contig of 7792 bp in length
 * 21211 22210: gap of 1000 bp
 * 22211 116154: contig of 93944 bp in length
 * 116155 117154: gap of 1000 bp
 * 117155 142308: contig of 25154 bp in length
 * 142309 143308: gap of 1000 bp
 * 143309 164162: contig of 20854 bp in length.

FEATURES

source

1. 164162
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone_lib="RPCT-11"
 /clone="R-98N22"
 BASE COUNT 52482 a 29753 c 28500 g 49421 t 4006 others
 ORIGIN

Query Match 100.0%; Score 15; DB 40; Length 164162;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

AC023555
 LOCUS 197540 bp DNA HTG 03-APR-2000
 DEFINITION Homo sapiens clone RP11-600K3, WORKING DRAFT SEQUENCE, 21 unordered
 pieces.
 ACCESSION AC023555
 VERSION AC023555.3 GI:7387357
 KEYWORDS HTG: HTGS_PHASE1, HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 (bases 1 to 197540)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 1 (bases 1 to 197540)
 Unpublished
 2 (bases 1 to 197540)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barne, N., Beda, F., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Burkett, G., Campoliano, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 DeRellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,
 Fenebor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Hagos, B., Hearford, A., Horton, L.,
 Grand-Pierre, N., Grant, G., Hagos, B., Jones, C., Kann, L., Karatas, A.,
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kamm, L., Karatas, A.,
 Klein, J., Landers, T., Largocque, K., Lehoczy, J., Levine, R.,
 Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
 McKean, P., McGuirk, A., McKernan, K., McSheeters, R., Meldrum, J.,
 Menais, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J., Naylor, J.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M.,
 Peterson, K., Pierre, N., Pisanl, C., Pollara, V., Raymond, C.,
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tittell, A.,
 Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
 Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and
 Zody, M.
 Direct Submission
 Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 3, 2000 this sequence version replaced gi:7249305.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: 600_K_3

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 184061 bases at least Q40
 Consensus quality: 190704 bases at least Q30
 Consensus quality: 193565 bases at least Q20
 Insert size: 200000; agarose-fp
 Insert size: 195540; sum-of-contigs
 Quality coverage: 4.3 in Q20 bases; agarose-fp
 Quality coverage: 4.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 21 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1308: contig of 1308 bp in length

```

* 1309 1408: gap of 100 bp
* 1409 3273: contig of 1865 bp in length
* 3274 3373: gap of 100 bp
* 3374 5999: contig of 2626 bp in length
* 6000 6099: gap of 100 bp
* 6100 8581: contig of 2482 bp in length
* 8582 8681: gap of 100 bp
* 8682 12149: contig of 3468 bp in length
* 12150 12249: gap of 100 bp
* 12250 15980: contig of 3731 bp in length
* 15981 16080: gap of 100 bp
* 16081 19563: contig of 3483 bp in length
* 19564 19663: gap of 100 bp
* 19664 22543: contig of 2880 bp in length
* 22544 22643: gap of 100 bp
* 22644 27530: contig of 4887 bp in length
* 27531 27630: gap of 100 bp
* 27631 33314: contig of 5684 bp in length
* 33315 33414: gap of 100 bp
* 33415 37745: contig of 4331 bp in length
* 37746 37845: gap of 100 bp
* 37846 46693: contig of 8848 bp in length
* 46694 46793: gap of 100 bp
* 46794 53246: contig of 6453 bp in length
* 53247 53346: gap of 100 bp
* 53347 66829: contig of 13483 bp in length
* 66830 66929: gap of 100 bp
* 66930 79204: contig of 12275 bp in length
* 79205 79304: gap of 100 bp
* 79305 92739: contig of 13435 bp in length
* 92740 92839: gap of 100 bp
* 92840 113251: contig of 20412 bp in length
* 113252 113351: gap of 100 bp
* 113352 132820: contig of 19469 bp in length
* 132821 132920: gap of 100 bp
* 132921 155098: contig of 22178 bp in length
* 155099 155198: gap of 100 bp
* 155199 173516: contig of 18318 bp in length
* 173517 173616: gap of 100 bp
* 173617 197540: contig of 23924 bp in length.

```

FEATURES

```

source
1. 197540
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="RPI1-600K3"
/clone_lib="RPI1-11 Human Male BAC"
1. 1308
misc_feature
/note="assembly_fragment"
1409..3273
misc_feature
/note="assembly_fragment"
3374..5999
misc_feature
/note="assembly_fragment"
6100..8581
misc_feature
/note="assembly_fragment"
8682..12149
misc_feature
/note="assembly_fragment"
12250..15980
misc_feature
/note="assembly_fragment"
16081..19563
misc_feature
/note="assembly_fragment"
19664..22543
misc_feature
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
22644..27530
misc_feature
/note="assembly_fragment"
27631..33314
misc_feature
/note="assembly_fragment"
33415..37745
misc_feature
/note="assembly_fragment"
clone_end:T7
vector_side:left"
37846..46693

```

```

misc_feature
/note="assembly_fragment"
46794..53246
misc_feature
/note="assembly_fragment"
53347..66829
misc_feature
/note="assembly_fragment"
66930..79204
misc_feature
/note="assembly_fragment"
79305..92739
misc_feature
/note="assembly_fragment"
92840..113251
misc_feature
/note="assembly_fragment"
113352..132820
misc_feature
/note="assembly_fragment"
132921..155098
misc_feature
/note="assembly_fragment"
155199..173516
misc_feature
/note="assembly_fragment"
173617..197540
misc_feature
/note="assembly_fragment"
BASE COUNT 60758 a 39052 c 39670 g 56053 t 2007 others
ORIGIN

```

Query Match 100.0%; Score 15; DB 72; Length 197540;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 caatagaagcattgc 15
|||||
Db 83840 CAATAGAGCATTGC 83854

```

RESULT 11

```

CEY1767
LOCUS CEY1767 207139 bp DNA HTG 03-DEC-1998
DEFINITION Caenorhabditis elegans chromosome II clone Y1767, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION Z96049
VERSION Z96049.1 GI:3378012
KEYWORDS HTG; HTGS_PHASE1.
SOURCE
ORGANISM

```

```

REFERENCE
AUTHORS Smye,R.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1998) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, UK and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jesesanger.ac.uk or twenematode.wustl.edu
On Aug 4, 1998 this sequence version replaced gi:2546913.
Order of segments is not known; 800 n's separate segments.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
be preserved.

```

COMMENT

```

Location/Qualifiers
1. 207139
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/chromosome="II"
/clone="Y1767"

```

FEATURES

```

source
BASE COUNT 68582 a 34583 c 35567 g 67606 t 801 others
ORIGIN

```

Query Match 100.0%; Score 15; DB 31; Length 207139;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatagaagccatgc 15
|||||
Db 97850 CAATAGAAGCCATGC 97864

RESULT 12
AL162574
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL162574 232489 bp DNA HTG 30-MAR-2000
Homo sapiens chromosome 13 clone RP11-282A11, *** SEQUENCING IN
PROGRESS ***, 48 unordered pieces.
AL162574
AL162574.2 GI:7378576
HTG: HTGS_PHASE1.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 232489)
Burton, J.
Submitted (30-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Mar 31, 2000 this sequence version replaced gi:7362707.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in
progress and the release of this data is based on the understanding
that the sequence may change as work continues. The sequence may
be contaminated with foreign sequence from E.coli, yeast, vector,
phage etc. Order of segments is not known; 800 n's separate
segments. Contig_ID: 00001 Length: 1241bp
Contig_ID: 00021 Length: 4239bp
Contig_ID: 00039 Length: 7392bp
Contig_ID: 00042 Length: 2426bp
Contig_ID: 00076 Length: 13588bp
Contig_ID: 00101 Length: 6905bp
Contig_ID: 00110 Length: 1013bp
Contig_ID: 00127 Length: 1325bp
Contig_ID: 00181 Length: 7108bp
Contig_ID: 00257 Length: 1486bp
Contig_ID: 00323 Length: 1933bp
Contig_ID: 00359 Length: 1039bp
Contig_ID: 00363 Length: 2084bp
Contig_ID: 00364 Length: 2463bp
Contig_ID: 00388 Length: 1306bp
Contig_ID: 00407 Length: 3117bp
Contig_ID: 00465 Length: 1074bp
Contig_ID: 00498 Length: 1346bp
Contig_ID: 00516 Length: 8917bp
Contig_ID: 00588 Length: 2925bp
Contig_ID: 00608 Length: 1404bp
Contig_ID: 00648 Length: 1466bp
Contig_ID: 00668 Length: 7176bp
Contig_ID: 00682 Length: 9578bp
Contig_ID: 00693 Length: 13363bp
Contig_ID: 00697 Length: 17250bp
Contig_ID: 00746 Length: 2116bp
Contig_ID: 00769 Length: 3454bp
Contig_ID: 00777 Length: 1915bp
Contig_ID: 00809 Length: 8622bp
Contig_ID: 00810 Length: 2574bp
Contig_ID: 00836 Length: 3282bp
Contig_ID: 00888 Length: 11506bp
Contig_ID: 00922 Length: 1074bp
Contig_ID: 00942 Length: 1009bp
Contig_ID: 00945 Length: 1684bp
Contig_ID: 01025 Length: 2522bp
Contig_ID: 01025 Length: 1337bp

Contig_ID: 01027 Length: 3586bp
Contig_ID: 01058 Length: 4573bp
Contig_ID: 01061 Length: 1524bp
Contig_ID: 01086 Length: 4972bp
Contig_ID: 01113 Length: 2372bp
Contig_ID: 01156 Length: 8035bp
Contig_ID: 01215 Length: 1023bp
Contig_ID: 01218 Length: 1740bp
Contig_ID: 01252 Length: 1422bp.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1241: contig of 1241 bp in length
* 1242 2041: gap of 800 bp
* 2042 6280: contig of 4239 bp in length
* 6281 7080: gap of 800 bp
* 7081 14472: contig of 7392 bp in length
* 14473 15272: gap of 800 bp
* 15273 17698: contig of 2426 bp in length
* 17699 18498: gap of 800 bp
* 18499 32066: contig of 13568 bp in length
* 32067 32866: gap of 800 bp
* 32867 39771: contig of 6905 bp in length
* 39772 40571: gap of 800 bp
* 40572 41584: contig of 1013 bp in length
* 41585 42384: gap of 800 bp
* 42385 43709: contig of 1325 bp in length
* 43710 44509: gap of 800 bp
* 44510 51617: contig of 7108 bp in length
* 51618 52417: gap of 800 bp
* 52418 53903: contig of 1486 bp in length
* 53904 54703: gap of 800 bp
* 54704 56636: contig of 1933 bp in length
* 56637 57436: gap of 800 bp
* 57437 58475: contig of 1039 bp in length
* 58476 59275: gap of 800 bp
* 59276 61359: contig of 2084 bp in length
* 61360 62159: gap of 800 bp
* 62160 64623: contig of 2464 bp in length
* 64624 65423: gap of 800 bp
* 65424 66723: contig of 1306 bp in length
* 66730 67529: gap of 800 bp
* 67530 70646: contig of 3117 bp in length
* 70647 71446: gap of 800 bp
* 71447 72578: contig of 1132 bp in length
* 72579 73378: gap of 800 bp
* 73379 74452: contig of 1074 bp in length
* 74453 75252: gap of 800 bp
* 75253 76798: contig of 1546 bp in length
* 76799 77598: gap of 800 bp
* 77599 86515: contig of 8917 bp in length
* 86516 87315: gap of 800 bp
* 87316 90240: contig of 2925 bp in length
* 90241 91040: gap of 800 bp
* 91041 92444: contig of 1404 bp in length
* 92445 93244: gap of 800 bp
* 93245 94710: contig of 1466 bp in length
* 94711 95510: gap of 800 bp
* 95511 102686: contig of 7176 bp in length
* 102687 103486: gap of 800 bp
* 103487 113064: contig of 9578 bp in length
* 113065 113864: gap of 800 bp
* 113865 127227: contig of 13363 bp in length
* 127228 128027: gap of 800 bp
* 128028 145277: contig of 17250 bp in length
* 145278 146077: gap of 800 bp
* 146078 148193: contig of 2116 bp in length
* 148194 148993: gap of 800 bp

* 148994 152447: contig of 3454 bp in length
* 152448 153247: gap of 800 bp
* 153248 155162: contig of 1915 bp in length
* 155163 155962: gap of 800 bp in length
* 155963 164654: contig of 8692 bp in length
* 164655 165454: gap of 800 bp
* 165455 168028: contig of 2574 bp in length
* 168029 168828: gap of 800 bp
* 168829 172110: contig of 3282 bp in length
* 172111 172910: gap of 800 bp
* 172911 184416: contig of 11506 bp in length
* 184417 185216: gap of 800 bp
* 185217 186290: contig of 1074 bp in length
* 186291 187090: gap of 800 bp
* 187091 188099: contig of 1009 bp in length
* 188100 188899: gap of 800 bp
* 188900 190583: contig of 1684 bp in length
* 190584 191383: gap of 800 bp
* 191384 193905: contig of 2522 bp in length
* 193906 194705: gap of 800 bp
* 194706 196042: contig of 1337 bp in length
* 196043 196842: gap of 800 bp
* 196843 200428: contig of 3586 bp in length
* 200429 201228: gap of 800 bp
* 201229 205801: contig of 4573 bp in length
* 205802 206601: gap of 800 bp
* 206602 208125: contig of 1524 bp in length
* 208126 208925: gap of 800 bp
* 208926 213897: contig of 4972 bp in length
* 213898 214697: gap of 800 bp
* 214698 217069: contig of 2372 bp in length
* 217070 217869: gap of 800 bp
* 217870 225904: contig of 8035 bp in length
* 225905 226704: gap of 800 bp
* 226705 227727: contig of 1023 bp in length
* 227728 228527: gap of 800 bp
* 228528 230267: contig of 1740 bp in length
* 230268 231067: gap of 800 bp
* 231068 232489: contig of 1422 bp in length.
Location/Qualifiers
1. 232489
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/clone_1lb="RP11-282A11"
/clone_1lb="RP11-11.1"

BASE COUNT 58033 a 39091 c 38984 g 58694 t 37687 others
ORIGIN

Query Match 100.0%; Score 15; DB 40; Length 232489;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatgc 15
|||||
Db 121982 CAATAGACCATGC 121996

RESULT 13
AF202305
LOCUS AF202305 1286 bp DNA PRI 14-DEC-1999
DEFINITION Homo sapiens beta-2 adrenergic receptor gene, complete cds.
ACCESSION AF202305
VERSION AF202305.1 GI:6573152
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1286)
Rupert, J.L. and Hochachka, P.W.
Beta-2-adrenergic receptor allele frequencies in two native

JOURNAL American populations
REFERENCE Unpublished
2 (bases 1 to 1286)
AUTHORS Rupert, J.L. and Hochachka, P.W.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1999) Zoology, University of British Columbia,
6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
Location/Qualifiers

1. 1286
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q31-q33"
/note="Isolated from Quechua speaking Native American"
/product="beta-2 adrenergic receptor"
35. .1276
/codon_start=1
/product="beta-2 adrenergic receptor"
/protein_id="AAP17569.1"
/db_xref="GI:6573153"

BASE COUNT 282 a 347 c 334 g 320 t 3 others
ORIGIN
Query Match 97.3%; Score 14.6; DB 39; Length 1286;
Best Local Similarity 93.3%; Pred. No. 3.6e+02;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatgc 15
|||||
Db 76 CAATAGACCATGC 90

RESULT 14
BACIS2401 865 bp DNA BCT 26-APR-1993
LOCUS BACIS2401 865 bp DNA BCT 26-APR-1993
DEFINITION B. thuringiensis insertion element IS240-A protein gene, complete
ACCESSION M23740 J03315
VERSION M23740.1 GI:143110
KEYWORDS Insertion sequence.
SEGMENT 1 of 2
SOURCE B. thuringiensis israelensis (strain 402-72) DNA, clone PRX70.
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 865)
AUTHORS Delecluse, A., Bourgoin, C., Klier, A. and Rapoport, C.
TITLE Nucleotide sequence and characterization of a new insertion
JOURNAL element, IS240 from Bacillus thuringiensis israelensis
MEDLINE Plasmid 21, 71-78 (1989)
COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by A. Delecluse, 04-APR-1989.
FEATURES
source Location/Qualifiers
1. 865
/organism="Bacillus thuringiensis"
/db_xref="taxon:1428"
repet_region 2. 16
CDS /note="left inverted repeat"
94. .801
/note="IS240-A protein"
/codon_start=1
/transl_table=11

/protein_id="AAA22555.1"
/db_xref="GI:143113"

/translation="MEKENIFPKMKHYQADMLITWYRWYLRYNLSPFDLVEEMERGLS
LSHTTMRWVHOYGPENLNERIKHLKRTNDSRMDPTYIKKIGENMYLYRAVDSGNT
LDFYLSKKRDAAKACFLKALASEFHTVTPRVITVDGNKATYVAIRLKNESITPGM
PLRVKKRYLNMTEDDHRFTKKRIILNMLGKSMQTAVKMIAGIEAHMVKKGLKLRQ
SAONONRCIHQLFGLTA"

repeat_region

850..864
/note="right inverted repeat"

BASE COUNT 300 a 143 c 176 g 246 t
ORIGIN 53 bp upstream of Ball site.

Query Match

93.3%; Score 14; DB 1; Length 865;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 aatagaagcatgc 15
|||||

Db 696 AATAGAAGCATGC 709

RESULT 15

BACIS2402 865 bp DNA BCT 26-APR-1993
LOCUS B.thuringiensis insertion element IS240-B protein gene, complete
DEFINITION cds.

ACCESSION M23741.1 GI:143111

KEYWORDS Insertion sequence; transposon.

SEGMENT 2 of 2
B.thuringiensis israelensis (strain 402-72) DNA, clone PRX70.

SOURCE Bacillus thuringiensis
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/staphylococcus group; Bacillus.

REFERENCE

1 (bases 1 to 865)
Delecluse, A., Bourgouin, C., Klier, A. and Rapoport, G.

TITLE Nucleotide sequence and characterization of a new insertion
element, IS240 from Bacillus thuringiensis israelensis
JOURNAL Plasmid 21, 71-78 (1989)

MEDLINE 89265213

COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by A.Delecluse, 04-APR-1989.

FEATURES

source location/Qualifiers

1..865
/organism="Bacillus thuringiensis"

repeat_region

1..16
/note="left inverted repeat"

CDS

94..801
/note="IS240-B protein"

/codon_start=1

/transl_table=1

/protein_id="AAA22555.1"

/db_xref="GI:143114"

/translation="MEKENIFPKMKHYQADMLITWYRWYLRYNLSPFDLVEEMERGLS
LSHTTMRWVHOYGPENLNERIKHLKRTNDSRMDPTYIKKIGENMYLYRAVDSGNT
LDFYLSKKRDAAKACFLKALASEFHTVTPRVITVDGNKATYVAIRLKNESISYGM
PLRVKKRYLNMTEDDHRFTKKRIILNMLGKSMQTAVKMIAGIEAHMVKKGLKLRQ
SAONONRCIHQLFGLTA"

repeat_region

850..865
/note="right inverted repeat"

BASE COUNT 302 a 141 c 178 g 244 t
ORIGIN About 13.5 kb after segment 1; 53 bp upstream of Ball site.

Query Match

93.3%; Score 14; DB 1; Length 865;

Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 aatagaagcatgc 15
|||||

Db 696 AATAGAAGCATGC 709

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:07:20 ; Search time 148.16 Seconds
(without alignments)
25.330 Million cell updates/sec

Title: US-09-542-718-4
Perfect score: 15
Sequence: 1 caatagaagccatgc 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues
Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	15	100.0	1999 1	Beta-2 adrenalin r
2	15	100.0	3451 1	Human beta-2-adren
3	14	93.3	874 1	Bacillus thuringie
4	14	93.3	3940 1	Delta-endotoxin cr
5	14	93.3	4571 1	Delta-endotoxin cr
6	14	93.3	4954 1	Insecticidal (Dipt
7	13.4	89.3	1627 1	H. pylori GHPD 154
8	13.4	89.3	1707 1	Plasium sativum L p
9	13	86.7	116 1	Staphylococcus aut
10	13	86.7	417 1	DNA encoding a Sta
11	13	86.7	521 1	Stealth virus nucl
12	13	86.7	521 1	Stealth virus nucl
13	13	86.7	713 1	Cotton fibrous tis
14	13	86.7	740 1	Streptococcus pneu
15	13	86.7	1016 1	Staphylococcus aut
16	13	86.7	1115 1	Nucleotide sequenc
17	13	86.7	1627 1	H. pylori GHPD 154
18	13	86.7	3592 1	Rat receptor tyros
19	13	86.7	4212 1	Human prollyl-4-hyd
20	13	86.7	5012 1	Adzuki bean peroxi
21	13	86.7	9473 1	HIV-2 variant HIV-
22	13	86.7	13868 1	Staphylococcus aut
23	13	86.7	174484	Continuation (3 of
24	13	86.7	110000 1	Continuation (4 of
25	13	86.7	110000 1	Continuation (4 of
26	13	86.7	110000 1	Continuation (4 of
27	12.6	84.0	2274 1	Coding sequence fo
28	12.4	82.7	30 1	Drosophila decpen
29	12.4	82.7	33 1	Primer for amplify
30	12.4	82.7	54 1	Rabbit CERP hairpin
31	12.4	82.7	54 1	Human CERP hairpin
32	12.4	82.7	195 1	H. pylori secreted
33	12.4	82.7	213 1	Human genome fragm

C	34	12.4	82.7	264 1	V78913	Staphylococcus aut
C	35	12.4	82.7	286 1	O61129	Human brain Expres
C	36	12.4	82.7	777 1	V59681	Human secreted pro
C	37	12.4	82.7	799 1	X39753	Gastric cancer as
C	38	12.4	82.7	828 1	T68075	H. pylori secreted
C	39	12.4	82.7	883 1	X39951	Prostate cancer as
C	40	12.4	82.7	997 1	V59800	Human secreted pro
C	41	12.4	82.7	1047 1	T98702	Single strand bind
C	42	12.4	82.7	1269 1	V66948	Canine herpes viru
C	43	12.4	82.7	1515 1	X20215	Enterococcus faeca
C	44	12.4	82.7	1662 1	T88864	Internal PAD2-N DN
C	45	12.4	82.7	1727 1	V42976	Streptococcus pneu

ALIGNMENTS

RESULT	1	100.0%: Score 15; DB 1; Length 1999;
T93250	standard; CDNA to mRNA: 1999 BP.	
AC	T93250;	
DT	20-APR-1998 (first entry)	
DE	Beta-2 adrenalin receptor subtype coding sequence.	
KW	Beta-2 adrenalin subtype; cyanopindrol; agonist; antagonist;	
KM	asthmatic disease; ss.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	190..1431
FT		/*tag- a
PN	MO9735963-A1.	
PD	02-OCT-1997.	
PF	24-MAR-1997; J00982.	
PR	27-MAR-1996; JP-072914.	
PI	(DAIN) DAINIPPON PHARM CO LTD.	
PI	Fuji K, Furutani Y, Kawashima H, Nomura A, Yano K;	
DR	WPI: 97-489627/45.	
DR	P-PDB: W34320.	
PT	Novel beta-2 adrenalin receptor sub-type - useful for screening for	
PT	agonists and antagonists and researching asthmatic diseases	
PS	Disclosure; Page 27-30; 47pp; Japanese.	
CC	This sequence encodes the protein of the invention. The protein of the	
CC	invention is a beta-2 adrenalin receptor subtype with Kd value of	
CC	approximately 75 pM against 125I-cyanopindrol. The protein can be used in	
CC	screening for agonists and antagonists, which are useful in researching	
CC	asthmatic diseases.	
SQ	Sequence 1999 BP; 477 A; 513 C; 485 G; 524 T;	
Query Match	100.0%: Score 15; DB 1; Length 1999;	
Best Local Similarity	100.0%; Pred. No. 12;	
Matches 15; Conservative	0; Mismatches	0; Indels
Gaps	0;	
Qy	1 caatagaagccatgc 15	
Db	231 CAATAGAAGCCATGC 245	
RESULT	2	
V52614	V52614 standard; CDNA; 3451 BP.	
AC	V52614;	
ID	21-DEC-1998 (first entry)	
DE	Human beta-2-adrenergic receptor cDNA.	
KW	Beta-2-adrenergic receptor; human; asthma; beta-agonist;	
KM	polymorphism; ds.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	1588..2829
FT		/*tag- a
FT	variation	1633
FT		/*tag- b
FT		/note- "A to G substitution, results in Arg16
FT		to Gly amino acid change"

```

PN      MO8939477-A2.
PD      11-SEP-1998.
PF      26-FEB-1998: U03908.
PR      03-MAR-1997: US-811441.
PA      (BCHM ) BRIGHAM & WOMENS HOSPITAL.
PI      Boushey H, Chinchilli VM, Drazen JM, Fish JE, Ford JG,
PI      Martin RJ.
DR      WPI: 98-506372/43.
DR      P-PSDB: W75777.
PT      Diagnosing asthma patients predisposed to adverse beta-agonist
PT      reactions upon regular administration - by identifying patients
PT      homozygous for allele encoding Arg at position 16 of
PT      beta2-adrenergic receptor protein
PS      Disclosure; Page 33-35; 46pp; English.
CC      This cDNA sequence codes for human beta-2-adrenergic receptor (see
CC      W75777) having an arginine residue at position 16. A novel method
CC      for identifying individuals susceptible to adverse responses to
CC      regular administration of beta-agonists comprises: (a) identifying
CC      in a genomic nucleic acid sample from the individual first and
CC      second alleles of the beta 2-adrenergic receptor gene, and (b)
CC      classifying an individual as susceptible if first and second
CC      alleles both encode Arg at residue 16 of the beta 2-adrenergic
CC      receptor protein. Beta 2-adrenergic receptor gene alleles may be
CC      identified by any known method e.g. denaturing gel electrophoresis
CC      or PCR amplification (see also V52615-17). Identification
CC      or PCR amplification (see also V52615-17). Identification
CC      preferably comprises amplifying a portion of each allele which
CC      includes the sequence encoding residue 16, and optionally also
CC      comprises determining nucleotide sequences of these portions (e.g.
CC      by automated sequence analysis). The invention identifies a known
CC      polymorphism in the beta 2-adrenergic receptor gene as being linked
CC      to adverse responses to regular beta-agonist administration;
CC      position 16 of the encoded protein can be either Arg or Gly, and
CC      individuals homozygous for Arg16 are more susceptible.
CC      Sequence 3451 BP; 790 A; 873 C; 895 G; 893 T;
SQ

```

Query Match 100.0%; Score 15; DB 1; Length 3451;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      1 caataagaagccatgc 15
DB      1629 CATTAGAACCCATGC 1643

```

RESULT 3

```

ID      070084 standard; DNA; 874 BP.
AC      070084:
DT      04-APR-1995 (first entry)
DE      Bacillus thuringiensis transcriptional control region.
KW      transcriptional control sequence; expression; insecticidal protein;
KW      Bacillus thuringiensis israelensis; develop; host vector system;
KW      insecticidal strain; ds.
OS      Bacillus thuringiensis israelensis.
PN      J06205682-A.
PD      26-JUL-1994.
PF      25-FEB-1992: 037748.
PR      25-FEB-1992: JP-037748.
PA      (SUMO ) SUMITOMO CHEM CO LTD.
DR      WPI: 94-275520/34.
PT      Transcriptional control sequence - for the expression of Bacillus
PT      thuringiensis insecticidal proteins
PS      Claim 1; Page 7; 9pp; Japanese.
CC      070084 is a transcriptional control sequence isolated from Bacillus
CC      thuringiensis (Bt) var. israelensis. It can be used for the expression
CC      of Bacillus thuringiensis insecticidal proteins. The isolation of this
CC      sequence allows the development of a host vector system in Bacillus
CC      thuringiensis and creation of highly insecticidal strains. The new
CC      strains created in this way are not recombinants, the insecticidal
CC      proteins produced using Bt as a host can be applied in agriculture.
SQ      Sequence 874 BP; 334 A; 101 C; 161 G; 278 T;

```

```

OY      2 aatagaagccatgc 15
DB      23 AATAGAACCCATGC 36

```

Query Match 93.3%; Score 14; DB 1; Length 874;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4

```

ID      N93054 standard; DNA; 3940 BP.
AC      N93054:
DT      30-MAY-1990 (first entry)
DE      Delta-endotoxin crystal protein gene.
KW      Delta-endotoxin; crystal protein; insecticide; ss; PCC130;
KW      biological control agent.
OS      Bacillus thuringiensis subsp israelensis.
FH      Key location/Qualifiers
FT      cds 891
FT      /*tag= a
FT      /product=delta-endotoxin
FT      rds 879. .884
FT      /*tag= b
FT      /label=Shine-Dalgarno sequence

```

EP-296870-A.
 28-DEC-1988. 305772.
 24-JUN-1988: 05-067653.
 26-JUN-1987: US-067653.
 (DUPO) Du Pont de Nemours Co.
 Ellar DJ, Ward ES;
 WPI: 89-001322/01.
 P-PSDB: P94035.

DR DNA fragment encoding insecticidal protein - obtained from
 PT Bacillus thuringiensis subsp. israelensis, and used in
 PT microorganisms and plant cells.
 PS Disclosure; fig 5; 26pp; English.
 CC The nucleotide sequence is an insert in plasmid pCC130. The delta
 CC endotoxin protein is insecticidal and can be used to control insect pests
 CC esp. mosquitoes.
 CC See also N93059.
 CC Sequence 3940 BP; 1457 A; 603 C; 659 G; 1221 T;

Query Match 93.3%; Score 14; DB 1; Length 3940;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      2 aatagaagccatgc 15
DB      39 AATAGAACCCATGC 52

```

RESULT 5

```

ID      N93059 standard; DNA; 4571 BP.
AC      N93059:
DT      30-MAY-1990 (first entry)
DE      Delta-endotoxin crystal protein gene.
KW      Delta-endotoxin; crystal protein; insecticide; pCH130; ss;
KW      biological control agent.
OS      Bacillus thuringiensis subsp israelensis.
FH      Key location/Qualifiers
FT      misc_feature 891. .4430
FT      /*tag= a
FT      /product=delta-endotoxin crystal protein
FT      RBS 879. .884
FT      /*tag= b
FT      /label= Shine-Dalgarno sequence.

```

EP-296870-A.
 28-DEC-1988. 305772.
 24-JUN-1988: 05-067653.

PR 26-JUN-1987: US-067653.
 PA (DUPO) Du Pont de Nemours CO.
 PI Ellar DJ, Ward ES;
 DR WPI: 89-001322/01.
 DR P-PSDB: P93715.
 PT DNA fragment encoding insecticidal protein - obtained from
 PT *Bacillus thuringiensis* subsp. *israelensis*, and used in
 PT microorganisms and plant cells.
 PS Disclosure: fig 7; 26pp: English.
 CC The sequence encodes the 130 kDa delta-endotoxin gene from
 CC *B. thuringiensis* subsp. *israelensis* and is a 4.46 kb insert of pCH130.
 CC The delta endotoxin protein is insecticidal and can be used to control
 CC insect pests esp. mosquitoes.
 CC See also N93054.
 SQ Sequence 4571 BP: 1678 A: 665 C: 817 G: 1391 T:

Query Match 93.3%; Score 14; DB 1: Length 4571;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 aatagaagccatgc 15
 |||||||
 Db 39 AATAGAAGCCATGC 52

RESULT 6
 N81490
 ID N81490 standard; DNA: 4934 BP.
 AC N81490;
 DT 05-NOV-1990 (first entry)
 DE Insecticidal (Diptera) protein gene.
 KW Insecticidal protein gene; Diptera; Aedes; ss.
 OS *Bacillus thuringiensis israelensis*.
 FH Key Location/Qualifiers
 FT cds 1..4908
 FT /*tag= a
 FT /product=insecticidal protein
 FT /tag= b
 FT Intron 620..1368
 FT
 PN J63230090-A.
 PD 26-SEP-1988.
 PF 19-MAR-1987: 066844.
 PR 19-MAR-1987: JP-066844.
 PA (SUMO) Sumitomo Chem Ind KK.
 DR WPI: 88-311968/44.
 DR P-PSDB: P81034, P82589.
 PT New insecticidal protein of *Bacillus thuringiensis israelensis* strain
 PT - propd. by forming gene library from plasmid contg. insecticidal protein
 PT gene, and transforming host cell.
 PS Disclosure: 9pp: Japanese.
 CC A plasmid contg. the gene is isolated by forming a gene library
 CC from plasmid DNA of *B. thuringiensis israelensis* (HD 522 strain, USA
 CC Goldberg ONR60) followed by screening with anti-*israelensis* insecticidal
 CC protein Ig.
 CC The insecticidal protein is highly effective against Diptera, esp. Aedes.
 SQ Sequence 4934 BP: 1797 A: 749 C: 877 G: 1511 T:

Query Match 93.3%; Score 14; DB 1: Length 4934;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 aatagaagccatgc 15
 |||||||
 Db 517 AATAGAAGCCATGC 530

RESULT 7
 X14561/C
 ID X14561 standard; DNA: 1627 BP.
 AC X14561;
 DT 31-MAR-1999 (first entry)

DE H. pylori GHPO 1548 gene.
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KW peptic ulcer disease; ss.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT cds 21..1571
 FT /*tag= a
 FT
 PN W09043478-A1.
 PD 08-OCT-1998.
 PR 01-APR-1998: U06371.
 PR 29-JUL-1997: US-902615.
 PR 01-APR-1997: US-833457.
 PR 24-JUN-1997: US-881227.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (INRA-) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
 DR WPI: 98-542293/46.
 DR P-PSDB: W98842.
 PT New isolated Helicobacter polynucleotides - used to develop products
 PT for the diagnosis, prevention and treatment of Helicobacter
 PT infections and gastrointestinal diseases
 PS Claim 1: Page 1884-1887; 2054pp: English.
 CC This sequence represents a polynucleotide of the invention. It was
 CC isolated from Helicobacter pylori and encodes a H. pylori GHPO protein.
 CC The polypeptides can be used for preventing or treating Helicobacter
 CC infections, and gastroduodenal diseases associated with these
 CC infections, including acute, chronic, and atrophic gastritis, and peptic
 CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
 CC for the production of antibodies. The products can also be used for
 CC detection and diagnosis.
 SQ Sequence 1627 BP: 502 A: 286 C: 354 G: 485 T:

Query Match 89.3%; Score 13.4; DB 1: Length 1627;
 Best Local Similarity 93.3%; Pred. No. 95;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 caatagaagccatgc 15
 |||||||
 Db 69 CGATAGAAGCCATGC 55

RESULT 8
 Q033019/C
 ID Q033019 standard; CDNA: 1707 BP.
 AC Q033019;
 DT 07-MAY-1993 (first entry)
 DE Pisum sativum L PAL gene.
 KW Phenylalanine ammonia lyase; phenyl propanoid isoflavonoid; plant;
 KW pea; ss.
 OS Pisum sativum.
 PN J04330285-A.
 PD 18-NOV-1992.
 PF 26-APR-1991: 097697.
 PR 26-APR-1991: JP-097697.
 PA (TAKS-) TAKASAGO PERFUMERY CO LTD.
 DR WPI: 93-003497/01.
 PT New phenylalanine ammonia lyase gene from *Pisum sativum* L -
 PT plasmid contg. gene, and E. coli transformed with plasmid, for
 PT prepn. of lyase.
 PS Disclosure: Page 8; 9pp: Japanese.
 CC The sequence is that of the phenylalanine ammonia lyase (PAL) gene
 CC from *Pisum sativum* L. The gene may be used to produce PAL
 CC recombinantly. Plants contg. the PAL gene can contain a large amt.
 CC of phenyl propanoid isoflavonoid.
 CC See also Q31985.
 SQ Sequence 1707 BP: 496 A: 345 C: 377 G: 489 T:

Query Match 89.3%; Score 13.4; DB 1: Length 1707;
 Best Local Similarity 93.3%; Pred. No. 96;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 caatagaagcatgc 15
 DB 182 CTAATGAGGCATGC 168

RESULT 9

V75746
 ID V75746 standard; DNA; 116 BP.
 AC V75746;
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #1435.
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 PN EP-786519-A2.
 PD 30-JUL-1997.
 PF 07-JAN-1997: 100117.
 PR 05-JAN-1996: US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA;
 DR WPI: 97-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 PS Claim 1; Page 1956; 3271pp; English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 of the invention. The DNA sequences are recorded on a computer readable
 medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 116 BP; 38 A; 21 C; 14 G; 42 T;

Query Match 86.7%; Score 13; DB 1; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 aatagaagcatgc 14
 DB 14 AATGAGGCATGC 26

RESULT 10
 T83720
 ID T83720 standard; DNA; 417 BP.
 AC T83720;
 DT 16-JUL-1998 (first entry)
 DE DNA encoding a Staphylococcus aureus protein of unknown function.
 KW Staphylococcus aureus protein; ribozyme; antisense sequence; control;
 KW Staphylococcal gene; regulatory element; bacterial gene expression;
 KW vaccine; Staphylococcal infection; food poisoning; scalded skin syndrome;
 KW toxic shock syndrome; ss.
 OS Staphylococcus aureus.
 FH Key Location/Qualifiers
 FT CDS complement(44..415)
 PN /*tag= a
 MO9730070-A1.

PD 21-AUG-1997.
 PF 19-FEB-1997: U02318.
 PR 20-FEB-1996: US-011888.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO,
 PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
 DR WPI: 97-424969/39.
 DR P-PSDB: W27751.

PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
 PT to isolate antimicrobial compounds, and in vaccines against S.
 PT aureus infection
 PS Claim 9; Page 638; 989pp; English.
 CC The present sequence encodes a Staphylococcus aureus protein of
 CC unknown function. The present sequence was isolated from a
 CC library of clones of S. aureus WCUH 29 in Escherichia coli. The DNA
 CC sequence can be used in the construction of ribozymes and antisense
 CC sequences to control the expression of Staphylococcal genes. The DNA
 CC sequence is also useful as a source of regulatory elements for the
 CC control of bacterial gene expression. The encoded protein may be used
 CC to produce vaccines to enable a host to produce specific antibodies
 CC with antibacterial action. These vaccines and antibodies would protect
 CC a host against invasion by S. aureus, and conditions relating to
 CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scalded
 CC skin syndrome, and toxic shock syndrome.
 SQ Sequence 417 BP; 134 A; 66 C; 78 G; 132 T;

Query Match 86.7%; Score 13; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatagaagcat 13
 DB 376 CAAATGAGGCAT 388

RESULT 11

V10144
 ID V10144 standard; DNA; 521 BP.
 AC V10144;
 DT 29-MAY-1998 (first entry)
 DE Stealth virus nucleic acid in plasmid #4.
 KW Stealth virus; chronic fatigue syndrome; CFS; disease; detection;
 KW medical diagnostic; veterinary diagnostic; agricultural diagnostic;
 KW quality control; ss.
 OS Stealth virus.
 PN US5703221-A.
 PM 30-DEC-1997.
 PD 05-JUN-1995: 463115.
 PF 05-JUN-1995: US-463115.
 PR 23-MAY-1991: US-704814.
 PR 20-SEP-1991: US-763039.
 PR 22-MAY-1992: US-887502.
 PR 23-NOV-1993: US-157811.
 PA (MART//) MARTIN W J.
 PI Martin WJ;
 DR WPI: 98-076485/07.

PT Stealth virus nucleic acid molecule - useful to detecting stealth
 PT virus, e.g. in chronic fatigue syndrome diagnosis
 PS Example 5; Fig 1; 82pp; English.
 CC V10141-V10210 represent nucleic acid sequences obtained from a
 CC patient with a stealth virus infection, namely chronic fatigue
 CC syndrome (CFS). Such nucleic acid sequences can be used to detect
 CC the stealth virus in medical, veterinary and agricultural diagnostics
 CC and in industrial and pharmaceutical biological quality control, e.g.
 CC to diagnose a disease associated with the stealth virus.
 SQ Sequence 521 BP; 119 A; 146 C; 117 G; 105 T;

Query Match 86.7%; Score 13; DB 1; Length 521;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 caatagaagccat 13
 |||||
 Db 218 CAATAGAAGCCAT 230

RESULT 12

ID V11957
 ID V11957 standard: DNA; 521 BP.

AC V11957;

DE 14-AUG-1998 (first entry)

DE Stealth virus plasmid 4 DNA.

KM Chronic fatigue syndrome; CFS; vaccine; cytopathic effect;

KM detection; ss.

OS Stealth virus.

PN USS753448-A.

PD 19-MAY-1998.

PE 05-JUN-1995; 465388.

PR 23-MAY-1991; US-704814.

PR 20-SEP-1991; US-763039.

PR 22-MAY-1992; US-887502.

PR 23-NOV-1993; US-157811.

PA (MART) MARTIN W J.

PI Martin WJ;

PI WPI: 98-311405/27.

PT Stealth virus contained in MRC-5 cell line, ATCC number VR2343 - 1s

PT useful as a vaccine against chronic fatigue syndrome

PS Disclosure; Fig 1A; 9pp; English.

CC V11954-V12024 are plasmids which contain fragments of a stealth virus

isolated from a patient, D.W. Such stealth virus fragments can be used

as vaccine against chronic fatigue syndrome (CFS). This illness causes

unexplained fatigue lasting more than 6 months and greater than 508

reduction in an infected persons normal level of activity. The virus

causes a cytopathic effect (CPE) to fibroblast cells observed in

CC culture, characterised by the appearance of rounded, slightly enlarged,

CC refractile cells in the culture.

SO Sequence 521 BP; 119 A; 146 C; 117 G; 105 T;

Query Match 86.7%; Score 13; DB 1; Length 521;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 caatagaagccat 13
 |||||

Db 218 CAATAGAAGCCAT 230

RESULT 13

ID T60570/c
 ID T60570 standard: cDNA to mRNA; 713 BP.

AC T60570;

DT 19-JUN-1997 (first entry)

DE Cotton fibrous tissue specific gene Gh2.

KM Fibrous tissue: cotton fibre; cotton flower; Gossypium barbadense;

KM Gossypium hirsutum; ss.

OS Gossypium hirsutum.

PH Key Location/Qualifiers

FT cds 37..447

FT /*tag- a

FT /product- fibrous tissue specific protein

PN J09075093-A.

PD 25-MAR-1997.

PE 20-FEB-1996; 031987.

PR 21-FEB-1995; US-391966.

PR 29-DEC-1995; US-580545.

PA (TOYM) TOYBO KK.

PA (UYTE-) UNIV TEXAS TECH.

DR WPI: 97-239274/22.

DR P-PSDB: W15762.

PT Cotton fibrous tissue gene - used to produce transformants with

PT Improved cotton fibres, and improved yield

PS Claim 2; Page 22-23; 37pp; Japanese.

CC T60567-T60571 represent cotton fibrous tissue genes of the invention.
 CC These sequences are specifically expressed during the growth of cotton
 CC fibre. The genes were isolated from the polyA RNA of the ovules of a
 CC cotton flower of Gossypium barbadense, and G. hirsutum. These sequences,
 CC and vectors containing them are used in the preparation of cotton
 CC transformants, with improved cotton fibres and improved yield.
 SO Sequence 713 BP; 200 A; 131 C; 161 G; 221 T;

Query Match 86.7%; Score 13; DB 1; Length 713;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 caatagaagccat 13
 |||||

Db 49 CAATAGAAGCCAT 37

RESULT 14

ID V37345/c
 ID V37345 standard: DNA; 740 BP.

AC V37345;

DT 13-OCT-1998 (first entry)

DE Streptococcus pneumoniae coding region.

KM coding region; ORF: open reading frame; antibacterial;

KM infection; prevention; meningitis; ss.

OS Streptococcus pneumoniae.

PH Key Location/Qualifiers

FT cds complement (1..261)

FT /*tag- a

FT /product- unknown

PN W09819688-A1.

PD 14-MAY-1998.

PE 27-OCT-1997; U19226.

PR 01-NOV-1996; US-029930.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,

PI Reid RH, Zarfos PN;

DR WPI: 98-286586/25.

DR P-PSDB: W60945.

PT New isolated nucleic acids from Streptococcus pneumoniae - useful,

PT e.g. for identifying anti-bacterial(s) for treatment and prevention

PT of meningitis

PS Claim 1; Page 47; 130pp; English.

CC The sequence is that of a coding region isolated from

CC S. pneumoniae. Its encoded protein, or agonists of it,

CC may be useful as an antibacterial for treatment or

CC prevention of infection, specifically caused by S. pneumoniae

CC (particularly meningitis) but possibly also Helicobacter

CC pylori (ulcers and gastric cancer). It may be of particular

CC use before insertion of an in-dwelling device or any other

CC invasive procedure. The protein, or nucleic acid encoding

CC it, can also be used in vaccines to induce a cellular

CC and/or humoral immune response, or to screen for other

CC antibacterials. The DNA may also contain flanking sequences

CC that are potential sources of control elements for bacterial

CC gene expression. Detecting a sequence encoding the protein

CC can be used diagnostically, e.g. to detect a mutation for

CC serotyping or classifying infectious agents.

SO Sequence 740 BP; 266 A; 141 C; 160 G; 173 T;

Query Match 86.7%; Score 13; DB 1; Length 740;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 caatagaagccat 13
 |||||

Db 356 CAATAGAAGCCAT 344

RESULT 15

V75239/C
 ID V75239 standard: DNA; 1016 BP.
 AC V75239;
 DT 16-MAR-1999 (first entry)
 DE Staphylococcus aureus contig SEQ ID #928.
 KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 PN EP-786519-A2.
 PD 30-JUL-1997.
 PR 07-JAN-1997; 100117.
 PR 05-JAN-1996; US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA;
 DR WPI: 97-374922/35.
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 PS Claim 1; Page 1751; 3271pp; English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, scalded skin syndrome, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 1016 BP; 344 A; 158 C; 151 G; 359 T;

Query Match 86.7%; Score 13; DB 1; Length 1016;
 Best Local Similarity 100.0%; Pred.No. 1.5e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 atagaagccatgc 15
 ||||||||||||
 DB 750 ATGAGAGCCATGC 738

Search completed: September 12, 2000, 23:07:23
 Job time: 4105 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:04:36 ; Search time 112.94 Seconds
(without alignments)
18.269 Million cell updates/sec

Title: US-09-542-718-4

Perfect score: 15

Sequence: 1 catcagaagcagtc 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 1000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA:*
1: /cgn2_6/ptodata/1/lna/3A_COMB.seq:*
2: /cgn2_6/ptodata/1/lna/3C_COMB.seq:*
3: /cgn2_6/ptodata/1/lna/3D_COMB.seq:*
4: /cgn2_6/ptodata/1/lna/3D_COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6_COMB.seq:*
6: /cgn2_6/ptodata/1/lna/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	1242	6	PCT-US91-00909-3
2	14	93.3	2121	4	US-08-633-879C-17
3	13.4	89.3	737	4	US-08-963-743-7
4	13	86.7	39	1	US-08-095-726-19
5	13	86.7	39	1	US-08-096-043-16
6	13	86.7	39	1	US-08-093-577-12
7	13	86.7	39	1	US-08-096-623A-24
8	13	86.7	521	2	US-08-463-115-10
9	13	86.7	521	2	US-08-465-388-10
10	13	86.7	713	4	US-08-580-545B-7
11	13	86.7	1863	5	US-09-126-646-1
12	13	86.7	2121	4	US-08-633-879C-17
13	13	86.7	3592	3	US-08-469-537A-100
14	12.4	82.7	33	5	US-08-836-402B-1
15	12.4	82.7	54	2	US-08-363-240A-1055
16	12.4	82.7	54	2	US-08-363-240A-1141
17	12.4	82.7	948	4	US-09-092-025A-1
18	12.4	82.7	1133	5	US-08-811-177A-3
19	12.4	82.7	1269	5	US-08-680-726A-67
20	12.4	82.7	1662	5	US-08-811-177A-1
21	12.4	82.7	2044	2	US-08-680-726A-63
22	12.4	82.7	2044	2	US-08-680-726A-64
23	12.4	82.7	2359	1	US-08-188-582-4
24	12.4	82.7	2359	1	US-08-646-715-4
25	12.4	82.7	2728	5	US-08-836-402B-7
26	12.4	82.7	4060	2	US-08-164-292B-1

27	12.4	82.7	4060	2	US-08-164-292B-3	Sequence 3, Appl1
28	12.4	82.7	4060	2	US-08-164-292B-5	Sequence 5, Appl1
29	12.4	82.7	4060	2	US-08-164-292B-7	Sequence 7, Appl1
30	12.4	82.7	4060	5	US-08-845-623-1	Sequence 1, Appl1
31	12.4	82.7	4060	5	US-08-845-623-3	Sequence 3, Appl1
32	12.4	82.7	4060	5	US-08-845-623-5	Sequence 5, Appl1
33	12.4	82.7	4060	5	US-08-845-623-7	Sequence 7, Appl1
34	12.4	82.7	4060	5	US-08-212-133A-7	Sequence 7, Appl1
35	12.4	82.7	7493	2	US-08-474-503-5	Sequence 5, Appl1
36	12.4	82.7	7493	3	US-08-670-707A-5	Sequence 5, Appl1
37	12.4	82.7	7493	6	PCT-US94-13200-5	Sequence 5, Appl1
38	12.4	82.7	12047	4	US-09-022-461-1	Sequence 1, Appl1
39	12.4	82.7	15328	4	US-08-888-497-33	Sequence 33, Appl1
40	12.4	82.7	15328	6	PCT-US94-07926-33	Sequence 33, Appl1
41	12	80.0	40	1	US-08-443-957-18	Sequence 18, Appl1
42	12	80.0	68	5	US-08-776-900C-40	Sequence 40, Appl1
43	12	80.0	450	1	US-08-090-523-28	Sequence 28, Appl1
44	12	80.0	450	1	US-08-398-627-28	Sequence 28, Appl1
45	12	80.0	450	2	US-08-596-024-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
PCT-US91-00909-3
Sequence 3, Application PC/TUS9100909
GENERAL INFORMATION:
APPLICANT: Sledziwski, Andrzej Z.
TITLE OF INVENTION: Methods of Producing Hybrid G
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: United States
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
PCT-US91/00909
CLASSIFICATION: 435
FILING DATE: 19910208
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 990008.408PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
TELEX: 3723836
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1242 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1242
PCT-US91-00909-3
Query Match 100.0%; Score 15; DB 6; Length 1242;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 caatagaagccatgc 15
|||||
Db 42 CAATGACGCGCATGC 56

RESULT 2

US-08-633-879C-17
; Sequence 17, Application US/08633879C
; Patent No. 5928922
; GENERAL INFORMATION:
; APPLICANT: Kivirikko, Karl I.
; APPLICANT: Pihlajaniemi, Taina
; APPLICANT: Helakoski, Tarja I.
; APPLICANT: Annunen, Pia P.
; APPLICANT: Nissi, Riitta K.
; APPLICANT: No. 5928922elainen, Minna K.
; TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
; TITLE OF INVENTION: METHODS FOR PRODUCING THE SAME
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/633, 879C
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0041-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 96...272
; OTHER INFORMATION:
; US-08-633-879C-17

Query Match 93.3%; Score 14; DB 4; Length 2121;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 aatagaagccatgc 15
|||||
Db 1880 AATGACGCGCATGC 1893

RESULT 3
US-08-963-743-7
; Sequence 7, Application US/08963743

; Patent No. 5952548
; GENERAL INFORMATION:
; APPLICANT: JIN, Wei
; APPLICANT: HORNER, Harry T.
; APPLICANT: PALMER, Reid G.
; APPLICANT: SHOEMAKER, Randy C.
; TITLE OF INVENTION: Soybean Glucanases, Compounds Which
; TITLE OF INVENTION: Encode Therefor and Related Methods
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kristine H. Johnson
; STREET: 123 No. 5952548th College Avenue, Suite 219
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80524
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963, 743
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: JOHNSON, Kristine H.
; REGISTRATION NUMBER: 36,835
; REFERENCE/DOCKET NUMBER: P-1029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (970) 472-9650
; TELEFAX: (970) 472-9655
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-963-743-7

Query Match 89.3%; Score 13.4; DB 4; Length 737;
Best Local Similarity 93.3%; Pred. No. 55;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 caatagaagccatgc 15
|||||
Db 448 CAATGACGCGCATGC 462

RESULT 4
US-08-095-726-19
; Sequence 19, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; TITLE OF INVENTION: Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA

ZIP: 60680-0703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,726
FILING DATE: 21-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/785,566
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5530189val B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3128564972
TELEFAX: 3128567180
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-095-726-19

Query Match 86.7%; Score 13; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 aatagaagccatg 14
|||||
Db 9 AATAGAAGCCATG 21

RESULT 5
US-08-096-043-16
Sequence 16, Application US/08096043
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B
TITLE OF INVENTION: Lycopen Biosynthesis In
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amoco Corp., Patents and Licensing Dept
STREET: 200 E Randolph St
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60680-0703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,043
FILING DATE: 22-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/785,568
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5530189val B
TELECOMMUNICATION INFORMATION:

TELEPHONE: 3128567180
TELEFAX: 3128564972
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-096-043-16

Query Match 86.7%; Score 13; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 aatagaagccatg 14
|||||
Db 9 AATAGAAGCCATG 21

RESULT 6
US-08-093-577-12
Sequence 12, Application US/08093577
Patent No. 5545816
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B
TITLE OF INVENTION: Phytoene Biosynthesis In
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amoco Corp., Patents and Licensing Dept
STREET: 200 E Randolph St
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60680-0703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,577
FILING DATE: 19-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/785,569
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5545816val B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3128567180
TELEFAX: 3128564972
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-093-577-12

Query Match 86.7%; Score 13; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 aatagaagccatg 14
|||||
Db 9 AATAGAAGCCATG 21

RESULT 7
US-08-096-623A-24
; Sequence 24, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indranil
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Hwei-Che B.
; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Welsh & Katz, Ltd.
; STREET: 120 S. Riverside Plaza, 22nd Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,623A
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,061
; FILING DATE: 09-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/662,921
; FILING DATE: 28-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/562,674
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/525,551
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,613
; FILING DATE: 02-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gansson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: AMO-006.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 655-1500
; TELEFAX: (312) 655-1501
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-096-623A-24

Query Match 86.7%; Score 13; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 aatagaagccatg 14
|||||

Db 9 AATAGAAGCCATG 21

RESULT 8
US-08-463-115-10
; Sequence 10, Application US/08463115
; Patent No. 5703221
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; TITLE OF INVENTION: AND RELATED VACCINES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Fastseq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,115
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: including application
; APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: No. 5703221, 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION:
; US-08-463-115-10

Query Match 86.7%; Score 13; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatagaagccat 13
|||||

Db 218 CAATAGAAGCCAT 230
RESULT 9
US-08-465-388-10
; Sequence 10, Application US/08465388
; Patent No. 5753488

GENERAL INFORMATION:
APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
TITLE OF INVENTION: AND RELATED VACCINES
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatilble
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,388
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/157,811
FILING DATE: No. 5753488ember 23, 1993
APPLICATION NUMBER: 07/887,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 213/300
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 521 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
OTHER INFORMATION:
US-08-465-388-10

Query Match 86.7%; Score 13; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccat 13
|||||
DB 218 CAATAGAGCCAT 230

RESULT 10
US-08-580-545B-7/c
Sequence 7, Application US/08580545B
Patent No. 5932713
GENERAL INFORMATION:
APPLICANT: Yoshitaka, Kasukabe
APPLICANT: Koichi, Fujisawa
APPLICANT: Susumu, Nishiguchi
APPLICANT: Yoshitaka, Mekawa
APPLICANT: Randy, Allen
TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fien & Richardson P.C.
STREET: 601 Thirteenth Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatilble
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580,545B
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bretschneider, Barry E.
REGISTRATION NUMBER: 28,055
REFERENCE/DOCKET NUMBER: 04473/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/783-2331
TELEFAX: 202/783-5070
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-580-545B-7

Query Match 86.7%; Score 13; DB 4; Length 713;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccat 13
|||||
DB 49 CAATAGAGCCAT 37

RESULT 11
US-09-126-646-1/c
Sequence 1, Application US/09126646
Patent No. 6001623
GENERAL INFORMATION:
APPLICANT: BRUN, KIMBERLY A.
APPLICANT: CREASY, CARETHA L.
TITLE OF INVENTION: DUNNINGTON, DAMIEN J.
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Ratner & Prestila
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatilble
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,646
FILING DATE: 31-JUL-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-126-646-1

Query Match 86.7%; Score 13; DB 5; Length 1863;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 atagaagccatgc 15
|||||
Db 576 ATAGAGCCATGC 564

RESULT 12
US-08-633-879C-17/C
Sequence 17, Application US/08633879C
Patent No. 5928922
GENERAL INFORMATION:
APPLICANT: Kivirikko, Karl I.
APPLICANT: Pihlajaniemi, Taina
APPLICANT: HeLaakoski, Tarja I.
APPLICANT: Annunen, Pia P.
APPLICANT: Nissi, Riitta K.
APPLICANT: No. 5928922elaine, Minna K.
TITLE OF INVENTION: '2 SUBUNIT OF PROLYL-4-HYDROXYLASE
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING SUCH SUBUNIT AND
TITLE OF INVENTION: METHODS FOR PRODUCING THE SAME
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM Compatible
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/633,879C
FILING DATE: 10-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0041-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2121 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 96...272
OTHER INFORMATION:
US-08-633-879C-17

Query Match 86.7%; Score 13; DB 4; Length 2121;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 aatagaagccatg 14
|||||
Db 1946 AATGAGCCATG 1934

RESULT 13
US-08-469-537A-100
Sequence 100, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisongier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US89 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: US89 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: US89 07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Ph.D., Gall M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 3592 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 596..3444
NAME/KEY: modified_base
LOCATION: 56
OTHER INFORMATION: /mod_base= OTHER
OTHER INFORMATION: //label= N
OTHER INFORMATION: /note= "where N = G, A, C or T"

FEATURE:
NAME/KEY: modified_base
LOCATION: 3538
OTHER INFORMATION: /mod_base- OTHER
OTHER INFORMATION: /label- N
OTHER INFORMATION: /note- "Where N = G, A, C or T"
US-08-469-537A-100

Query Match 86.7%; Score 13; DB 3; Length 3592;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 caatagaagccatg 14
|||||
Db 1836 AATGAGACCATG 1848

RESULT 14
US-08-836-402B-1/C
Sequence 1, Application US/08836402B
Patent No. 6063988
GENERAL INFORMATION:
APPLICANT: Rudiger Hain, Regina Fischer
TITLE OF INVENTION: DNA SEQUENCE AND ITS USE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: Gateway 2000
OPERATING SYSTEM: DOS
SOFTWARE: Microsoft Windows 98
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,402B
FILING DATE: 02-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 44 40 200.7 (Germany)
FILING DATE: 10-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 9840-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
US-08-836-402B-1

Query Match 82.7%; Score 12.4; DB 5; Length 33;
Best Local Similarity 92.9%; Pred. No. 1.5e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
|||||
Db 26 CAATTGAACCATG 13

RESULT 15
US-08-363-240A-1055
Sequence 1055, Application US/088363240A
Patent No. 5705388
GENERAL INFORMATION:
APPLICANT: Couture, Larry
APPLICANT: McSwiggen, James
APPLICANT: Bisgaler, Charles
APPLICANT: Page, Michael
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: PREVENTION, INHIBITION OF
TITLE OF INVENTION: PROGRESSION AND REGRESSION
TITLE OF INVENTION: OF VASCULAR DISEASES
NUMBER OF SEQUENCES: 1243
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,240A
FILING DATE: December 23, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 210/096
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1055:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-363-240A-1055

Query Match 82.7%; Score 12.4; DB 2; Length 54;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
|||||
Db 5 CAAAGAAGCCCAUG 18

Search completed: September 12, 2000, 23:04:39
Job time: 3946 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 22:46:49 : Search time 1893.64 Seconds
(without alignments)
34.931 Million cell updates/sec

Title: US-09-542-718-4

Perfect score: 15

Sequence: 1 caatagaagcctatgc 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_gss1:*
94: gb_gss2:*
95: gb_gss3:*
96: gb_gss4:*
97: em_gss1:*
98: em_gss2:*
99: em_gss3:*
100: em_gss4:*
101: gb_gss5:*
102: gb_gss6:*
103: gb_gss7:*
104: gb_gss8:*
105: gb_gss9:*
106: em_gss5:*
107: em_gss6:*
108: em_gss7:*
109: em_gss8:*
110: em_gss9:*
111: em_gss10:*
112: em_gss11:*
113: gb_gss10:*
114: gb_gss11:*
115: em_gss12:*
116: gb_gss12:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	15	100.0	287	74	AM598789	g88804.y
2	14	93.3	219	50	AV132414	AV132414
3	14	93.3	312	104	AQ553837	AQ553837
4	14	93.3	398	118	AZ008470	AZ008470
5	14	93.3	409	96	AQ317207	AQ317207
6	14	93.3	435	95	AQ195952	AQ195952
7	14	93.3	471	113	AQ678011	AQ678011
8	14	93.3	490	114	AQ777734	AQ777734
9	14	93.3	515	79	AM674779	AM674779
10	14	93.3	515	104	AQ595678	AQ595678
11	14	93.3	519	94	AM561225	AM561225
12	14	93.3	539	94	AQ131615	AQ131615
13	14	93.3	560	38	A1314125	A1314125
14	14	93.3	561	118	AZ018937	AZ018937
15	14	93.3	618	42	A1659892	A1659892
16	14	93.3	659	103	AQ510816	AQ510816
17	13.4	89.3	173	62	AV368632	AV368632
18	13.4	89.3	192	93	AQ077922	AQ077922
19	13.4	89.3	197	21	AA104807	AA104807
20	13.4	89.3	197	60	AV322345	AV322345
21	13.4	89.3	202	35	A1115451	A1115451
22	13.4	89.3	203	49	AV085655	AV085655
23	13.4	89.3	209	73	AM501422	AM501422
24	13.4	89.3	233	27	AA511687	AA511687
25	13.4	89.3	233	27	AA869212	AA869212
26	13.4	89.3	235	59	AV241497	AV241497
27	13.4	89.3	238	101	AQ378683	AQ378683
28	13.4	89.3	239	118	AQ989917	AQ989917
29	13.4	89.3	242	30	AA717083	AA717083
30	13.4	89.3	261	49	AV067032	AV067032
31	13.4	89.3	268	63	AM037106	AM037106
32	13.4	89.3	271	93	AQ044058	AQ044058
33	13.4	89.3	292	64	AM166441	AM166441
34	13.4	89.3	297	120	B68841	B68841
35	13.4	89.3	299	23	AA241735	AA241735
36	13.4	89.3	317	73	AM488772	AM488772
37	13.4	89.3	337	31	AA832805	AA832805
38	13.4	89.3	342	91	W81900	W81900
39	13.4	89.3	344	21	AA129379	AA129379
40	13.4	89.3	354	72	AA457447	AA457447
41	13.4	89.3	362	79	AM654359	AM654359
42	13.4	89.3	368	95	AQ198499	AQ198499
43	13.4	89.3	373	40	A1468443	A1468443
44	13.4	89.3	376	23	AA258617	AA258617
45	13.4	89.3	376	91	W29318	W29318

ALIGNMENTS

RESULT 1
LOCUS AM598789 287 bp mRNA
DEFINITION g88804.y1 Moss EST library PPU Physcomitrella patens cDNA clone
PEP_SOURCE_ID: PPU161308 5', mRNA sequence.

ACCESSION AM598789
VERSION AM598789.1 GI:7286302
KEYWORDS EST.
SOURCE Physcomitrella patens.
ORGANISM Physcomitrella patens
REFERENCE Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 287)
AUTHORS Quatrano, R., Bashardes, S., Cove, D., Cunliffe, A., Knight, C., Clifton, S., Maria, M., Hillier, L., Pape, D., Martin, J., Wyllie, T., Underwood, K., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Steptoe, M., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
Leeds/Wash u Moss Est Project
Unpublished (1999)
On Jul 8, 1999 this sequence version replaced gi:5422558.
COMMENT Contact: Ralph Quatrano
Leeds/Wash u Moss Est Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Libraries were constructed by Dr. Stavros Bashardes as part of the Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and Washington Univ. in St. Louis (USA) DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
Seq primer: -40RP from Gibco.

FEATURES

source
1..287
/organism="Physcomitrella patens"
/db_xref="taxon:3218"
/clone="PEP_SOURCE_ID: PPU161308"
/clone_lib="Moss EST library PPU"
/tissue_type="protonemata: 7 day old tissue
ammonium-grown."
/lab_host="DH10B"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Construction of the cDNA library was carried out using Stratagene's 'UniZAP' cDNA synthesis kit. cDNA was constructed using an oligo dt primer/linker that contains a XhoI site within it. Following ds cDNA synthesis, EcoRI adapters were ligated to the blunt ends and sample was digested with XhoI. The result is cDNA with an EcoRI sticky end on one side and a XhoI sticky end on the other. This cDNA was ligated directionally in UniZAP arms. The vector is designed containing the Bluescript sequence as well as lambda DNA and cDNA is cloned within this plasmid sequence. The vector was then packaged using Gold digpackaging extracts. Library was grown in XLBlue MRF' cells and amplified. The library was excised by mass excision using Stratagene's 'Mass excision kit' that uses exsist as a helper phase that releases the plasmid sequence and circularises it as single stranded plasmids that are then packaged (by helper phage) and secreted out of the host cell as phagelids. SOLR cells were transformed with phagelids and the library was plated out on LB-amp plates to select for transformants. Approximately 1,000,000 colonies were grown and recovered. The double stranded plasmid library was recovered by using Qiaagen Midi prep kit. 2 micro grams of each library were used to transform DH10B cells by electroporation."

BASE COUNT 61 a 69 c 74 g 83 t
ORIGIN

Query Match 100.0%; Score 15; DB 74; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 caatagaagccatgc 15

```

Db      261 CAATAGAGCCATGC 275
|||||
RESULT  2
LOCUS   AV132414      219 bp      mRNA      EST      01-JUL-1999
DEFINITION
AV132414 Mus musculus C57BL/6J 11-day embryo Mus musculus cDNA
AV132414
VERSION AV132414.1 GI:5318649
KEYWORDS
SOURCE  house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
AUTHORS Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shireki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y.,
Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished (1999)
On Apr 7, 1998 this sequence version replaced gi:3035541.
COMMENT
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtrc.riken.go.jp
Thermosensitization and thermocytivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtrc.riken.go.jp) for
further details.
FEATURES
source
location/Qualifiers
1..219
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="2700089N23"
/clone_lib="Mus musculus C57BL/6J 11-day embryo"
/sex="mixed"
/dev_stage="11-day embryo"
BASE COUNT 78 a 59 c 50 g 28 t 4 others
ORIGIN
Query Match 93.3%; Score 14; DB 50; Length 219;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 aatagaagccatgc 15
|||||
Db 72 AATAGAGCCATGC 85
RESULT 3
LOCUS A0553837 312 bp DNA GSS 28-MAY-1999
DEFINITION RPCI-11-357N10.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-
357N10, genomic survey sequence.
ACCESSION A0553837
VERSION A0553837.1 GI:4913014
KEYWORDS GSS.

```

```

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT On Dec 15, 1999 this sequence version replaced gi:4213307.
Other GSSs: RPCI-11-357N10.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.bufileo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: Sp6
Class: BAC ends.
FEATURES
source
location/Qualifiers
1..312
/organism="Homo sapiens"
/db_xref="GDB:7637025"
/db_xref="taxon:9606"
/clone="RPCI-11-357N10"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site.1: EcoRI; Site.2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 110 a 60 c 82 g 60 t
ORIGIN
Query Match 93.3%; Score 14; DB 104; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 caatagaagccatgc 14
|||||
Db 88 CAATAGAGCCATGC 101
RESULT 4
LOCUS A2008470 398 bp DNA GSS 25-FEB-2000
DEFINITION RPCI-23-246M9.TJ RPCI-23 Mus musculus genomic clone RPCI-23-
246M9, genomic survey sequence.
ACCESSION A2008470
VERSION A2008470.1 GI:7083854
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., Megann, S., Tsegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT On Dec 15, 1999 this sequence version replaced gi:4574988.
Contact: Shaying Zhao
Department of Eukaryotic Genomics

```

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: szhac@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 246 row: M column: 9
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. .398
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-246M9"
/clone_id="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 112 a 70 c 70 g 135 t 11 others
ORIGIN

Query Match 93.3%; Score 14; DB 118; Length 398;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
|||||
Db 207 CAATGAGCCATG 194

RESULT 5
AQ317207 409 bp DNA GSS 04-MAY-1999
LOCUS RPCI11-106B3.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-106B3, genomic survey sequence.
DEFINITION
ACCESSION AQ317207

VERSION AQ317207.1 GI:4048458
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 409)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPCI11-106B3.TJ
Contact: Shaying Zhao, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from

Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. .409
/organism="Homo sapiens"
/db_xref="GDB:7540346"
/db_xref="taxon:9606"
/clone="RPCI-11-106B3"
/clone_id="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI; RPCI11 Human Male BAC Library"
BASE COUNT 142 a 85 c 112 g 70 t
ORIGIN

Query Match 93.3%; Score 14; DB 96; Length 409;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
|||||
Db 98 CAATGAGCCATG 111

RESULT 6
AQ199592 435 bp DNA GSS 20-APR-1999
LOCUS RPCI11-48H5.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-48H5, genomic survey sequence.
DEFINITION
ACCESSION AQ199592

VERSION AQ199592.1 GI:3611791
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 435)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPCI11-48H5.TJ
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. .435
/organism="Homo sapiens"
/db_xref="GDB:7518220"
/db_xref="taxon:9606"
/clone="RPCI-11-48H5"
/clone_id="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI; RPCI11 Human Male BAC Library"
BASE COUNT 126 a 84 c 95 g 130 t

ORIGIN

Query Match 93.3%; Score 14; DB 95; Length 435;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatgc 14
 |||||||
 Db 92 CAATAGAACCATGC 105

RESULT 7

LOCUS A0678011/c 471 bp DNA GSS 25-JUN-1999
 DEFINITION HS_5531.B2.G07-SP6 RPCI-11 Human Male BAC Library Homo
 sapiens genomic clone Plate-1107 Col-14 Row-N, genomic survey
 sequence.

ACCESSION A0678011

VERSION A0678011.1 GI:5226815

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 471)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,D., Young,D., Zhao,S., Adams,M.D. and
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pletier de Jong

(pleterdejong@med.buflalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buflalo.edu/ordering_bac.htm)

or from Research Genetics (info@resgen.com). BAC end Web Server:
 http://www.htsc.washington.edu

Plate: 1107 row: N column: 14

Seq primer: SP6

Class: BAC ends

High quality sequence stop: 471.

Location/Qualifiers

1. 471

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="plate-1107 Col-14 Row-N"

/clone_11b="RPCI-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI;
 Male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI Methylase. Size selected DNA was cloned into the
 pBAC3.6 vector at EcoRI sites"

BASE COUNT 133 a 81 c 93 g 157 t 7 others

ORIGIN

Query Match 93.3%; Score 14; DB 113; Length 471;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 aatagaagccatgc 15
 |||||||

Db 455 AATAGAAGCCATGC 442

RESULT 8
 LOCUS A0777734 490 bp DNA GSS 02-AUG-1999
 DEFINITION HS_2190.B2.A07_MR CIT Approved Human Genomic Sperm Library D
 Homo sapiens genomic clone Plate-2190 Col-14 Row-B, genomic
 survey sequence.

ACCESSION A0777734

VERSION A0777734.1 GI:5680694

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 490)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,D., Young,D., Zhao,S., Adams,M.D. and
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT On Sep 10, 1998 this sequence version replaced gi:3554391.
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu

Plate: 2190 row: B column: 14

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 490.

Location/Qualifiers

1. 490

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="plate-2190 Col-14 Row-B"

/clone_11b="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: Sperm; Vector: pBelBAC11; BAC Clones in
 E-Coli DH10B"

BASE COUNT 142 a 85 c 96 g 166 t 1 others

ORIGIN

Query Match 93.3%; Score 14; DB 114; Length 490;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatgc 14
 |||||||

Db 109 CAATAGAAGCCATGC 122

RESULT 9

LOCUS AM674779 515 bp mRNA EST 11-APR-2000

DEFINITION ba59e06.y1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:2900866.5,
 similar to SW:SP49_HUMAN Q15427 SPLICOSOME ASSOCIATED PROTEIN 49
 ; mRNA sequence.

ACCESSION AM674779

VERSION AM674779.1 GI:7540089

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 515)
 AUTHORS NIH-MGC <http://www.ncbi.nlm.nih.gov/MGC/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT On Jul 7, 1999 this sequence version replaced gi:5407333.
 Other ESTs: ba59e06.x1
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov

TISSUE Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
imgl.gov/image/html/lresources.shtml
 Seq primer: -40RP from Gibco
 High quality sequence stop: 423.
 Location/Qualifiers

FEATURES
 source
 1..515
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2900866"
 /clone_lib="NIH-MGC_10"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies"
 BASE COUNT 137 a 117 c 136 g 124 t 1 others
 ORIGIN

Query Match 93.3%; Score 14; DB 79; Length 515;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 caatagaagccatgc 14
 |||||||
 Db 502 CAATGAGCCATGC 515

RESULT 10
 A0595678/c 515 bp DNA GSS 08-JUN-1999
 LOCUS HS_5459_B2.D06.T7A RPCR1-11 Human Male BAC Library Homo sapiens
 DEFINITION genomic clone Plate=1035 Col=12 Row=H, genomic survey sequence.
 ACCESSION A0595678
 VERSION A0595678.1 GI:5027264
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 515)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 515)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT On Dec 15, 1999 this sequence version replaced gi:4215181.
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887

Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCR1-11. For BAC library availability, please contact Pieter de Jong (piet@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (<http://www.hsc.washington.edu>) Plate: 1035 row: H column: 12
 Seq primer: 77
 Class: BAC ends
 High quality sequence stop: 515.
 Location/Qualifiers

FEATURES
 source
 1..515
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate-1035 Col-12 Row=H"
 /clone_lib="RPCR1-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"
 BASE COUNT 137 a 120 c 102 g 152 t 4 others
 ORIGIN

Query Match 93.3%; Score 14; DB 104; Length 515;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 aatagaagccatgc 15
 |||||||
 Db 90 AATGAGCCATGC 77

RESULT 11
 A0561225 519 bp mRNA EST 08-MAR-2000
 LOCUS 9a77b08.y1 Moss EST library PPU Physcomitrella patens cDNA clone
 DEFINITION pep_SOURCE_ID:PPU140116 5' similar to gb:gb1092085.1|nt092085 Arabidopsis thaliana clathrin assembly (PLANT); mRNA sequence.
 ACCESSION A0561225
 VERSION A0561225.1 GI:7207222
 KEYWORDS EST.
 SOURCE Physcomitrella patens.
 ORGANISM Physcomitrella patens.
 Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

REFERENCE 1 (bases 1 to 519)
 AUTHORS Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Underwood,K., Rheising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Steptoe,M., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R., Waterston,R. and Wilson,R.
 Leeds/Wash U Moss EST Project
 Unpublished (1999)
 On Jan 6, 2000 this sequence version replaced gi:6677556.
 Contact: Ralph Quatrano
 Leeds/Wash U Moss EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

TITLE Libraries are constructed by Dr. Stavros Bashlades as part of the Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and Washington Univ. in St. Louis (USA) DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 425.
 Location/Qualifiers

FEATURES

source

1. .519
 /organism="Physcomitrella patens"
 /db_xref="taxon:3218"
 /clone="PEP_SOURCE_ID:PPU140116"
 /clone_lib="Moss EST library PPU"
 /tissue_type="protonemata: 7 day old tissue
 ammonium-grown"
 /lab_host="DH10B"

/note="Vector: Bluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; Construction of the cDNA library was carried out
 using Strategenes 'Unizap - cDNA synthesis kit'. cDNA
 was constructed using an oligo dt primer/linker that
 contains a XhoI site within it. Following ds cDNA
 synthesis, EcoRI adapters were ligated to the blunt ends
 and sample was digested with XhoI. The result is cDNA
 with an EcoRI sticky end on one side and a XhoI sticky
 end on the other. This cDNA was ligated directionally in
 Unizap arms. The vector is designed containing the
 Bluescript sequence as well as lambda DNA and cDNA is
 cloned within this Bluescript sequence. The vector was
 then packaged using Gold gigaPackaging extracts. Library
 was grown in XLBlue MRF⁺ cells and amplified. The library
 was excised by mass excision using Strategenes 'Mass
 excision kit' that uses exassit as a helper phage that
 releases the Bluescript sequence and circularises it as
 single stranded plasmids that are then packaged (by helper
 phage) and secreted out of the host cell as phagemids.
 SOLR cells were transformed with phagemids and the library
 was plated out on LB-amp plates to select for
 transformants. Approximately 1,000,000 colonies were grown
 and recovered. The double stranded plasmid library was
 recovered by using Qulagen Midl prep kit. 2 micro grams of
 each library were used to transform DH10B cells by
 electroporation."

BASE COUNT 136 a 104 c 110 g 168 t 1 others
 ORIGIN

Query Match 93.38; Score 14; DB 74; Length 519;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
 ||||||||||||
 Db 395 CAATAGAAGCCATG 408

RESULT 12

LOCUS A0113165 539 bp DNA GSS 29-AUG-1998
 DEFINITION CIT-HSP-237509.TF CIT-HSP Homo sapiens genomic clone 237509,
 genomic survey sequence.
 ACCESSION A0113165

VERSION A0113165.1 GI:3489286
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 539)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building

JOURNAL Unpublished (1998)
 COMMENT Other_GSSs: CIT-HSP-237509.TF
 Contact: Mark Adams

Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200

Fax: 301 838 0208
 Email: mdadams@igf.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
http://www.igf.org/tdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.

FEATURES
 source location/Qualifiers

1. .539
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="237509"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBelOBAC11; Site.1: HindIII; Site.2:
 HindIII"

BASE COUNT 173 a 84 c 148 g 134 t
 ORIGIN

Query Match 93.38; Score 14; DB 94; Length 539;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
 ||||||||||||
 Db 503 CAATAGAAGCCATG 516

RESULT 13

LOCUS A1314125/c 560 bp mRNA EST 17-DEC-1998
 DEFINITION U12E03.X1 Sugano mouse kidney mklA Mus musculus cDNA clone
 IMAGE:1907836 3', mRNA sequence.
 ACCESSION A1314125
 VERSION A1314125.1 GI:4029368
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 560)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Geisel,S., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Schellenberg,K., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mcuseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:976032

Seq primer: custom primer used
 High quality sequence stop: 501.
 Location/Qualifiers
 1. .560

/organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1907836"
 /clone_lib="Sugano mouse kidney mklA"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pMT18s-FL3; Site.1: DraIII"

(CACTGCTG); Site 2: DraIII (CACCATGCTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TCGTGGCCCTACGCG], digested and cloned into distinct DraIII sites of the pMT18S-Fl3 vector (5' site CACTGCTG, 3' site CACCATGCTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCACCTCGACGACA.

BASE COUNT 180 a 160 c 71 g 149 t

ORIGIN

Query Match 93.3%; Score 14; DB 38; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 caatagaagccatg 14
|||||

DB 386 CAATAGAACCATG 373

RESULT 14
LOCUS AZ018937 561 bp DNA GSS 25-FEB-2000
DEFINITION RPCI-23-298P23, RV RPCI-23 Mus musculus genomic clone RPCI-23-298P23, genomic survey sequence.
ACCESSION AZ018937
VERSION AZ018937.1 GI:7094321
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 561)
Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatsman, S., Akirel, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.
Mouse BAC end Sequences from Library RPCI-23
Unpublished (1999)
On Dec 15, 1999 this sequence version replaced gi:4575467.
Other GSS: RPCI-23-298P23.FJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pletier de Jong (pleterdejong.med.bufileo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.bufileo.edu/orderingframe.htm>) or from Resea ch Genetics (infresgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 298 row: P column: 23
Seq primer: 17
Class: BAC ends.

FEATURES
source
1. 561
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-298P23"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies).

BASE COUNT 190 a 114 c 109 g 146 t 2 others

ORIGIN

Query Match 93.3%; Score 14; DB 119; Length 561;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 aatagaagccatgc 15
|||||

DB 472 AATAGAACCATGC 459

RESULT 15
LOCUS A1659892 618 bp mRNA EST 10-MAY-1999
DEFINITION t001803.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2249765 3' similar to gb:W94556 SINGLE-STRANDED DNA-BINDING PROTEIN MITOCHONDRIAL PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION A1659892
VERSION A1659892.1 GI:4763462
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 618)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
On Oct 8, 1998 this sequence version replaced gi:3728902.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmer-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.llnl.gov/dbp/image/image.html
Seq primer: -40UP from glibco
High quality sequence stop: 469.

FEATURES
source
1. 618
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2249765"
/clone_lib="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 985608-986759, 1101192-1101959, and 121928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 162 a 147 c 104 g 203 t 2 others

ORIGIN

Query Match 93.3%; Score 14; DB 42; Length 618;

Best Local Similarity 100.08; Pred. No. 5.6e+02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 caatagaagccatg 14
|||||
Db 562 CAATGAGGCCATG 549

Search completed: September 12, 2000, 22:46:55
Job time: 3797 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:03:21 ; Search time 930.45 Seconds
(without alignments)
24.935 Million cell updates/sec

Title: US-09-542-718-5
Perfect score: 13
Sequence: 1 cccaatgaagac 13

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 972840 seqs, 892348106 residues

Total number of hits satisfying chosen parameters: 1945680

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_om: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_pl1: *
8: gb_pl2: *
9: gb_pr1: *
10: gb_pr2: *
11: gb_pr3: *
12: gb_ro: *
13: gb_ste: *
14: gb_sy: *
15: gb_un: *
16: em_fun: *
17: em_hum1: *
18: em_hum2: *
19: em_in: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_ste: *
28: em_sy: *
29: em_un: *
30: em_v1: *
31: gb_htg1: *
32: gb_htg2: *
33: gb_in1: *
34: gb_in2: *
35: em_ba1: *
36: em_ba2: *
37: em_hum3: *
38: em_hum4: *
39: gb_pr4: *
40: gb_htg3: *
41: gb_htg4: *
42: gb_htg5: *
43: gb_htg6: *

44: gb_htg7: *
45: em_htg1: *
46: em_htg2: *
47: em_htg3: *
48: em_hum5: *
49: gb_pl3: *
50: gb_pr5: *
51: gb_htg8: *
52: gb_htg9: *
53: gb_htg10: *
54: gb_htg11: *
55: gb_htg12: *
56: gb_htg13: *
57: gb_htg14: *
58: gb_in3: *
59: gb_htg15: *
60: gb_htg16: *
61: gb_htg17: *
62: em_htg4: *
63: em_htg5: *
64: em_htg6: *
65: em_htg7: *
66: em_hum6: *
67: gb_htg18: *
68: gb_htg19: *
69: gb_htg20: *
70: gb_htg21: *
71: gb_htg22: *
72: gb_htg23: *
73: gb_htg24: *
74: gb_htg25: *
75: gb_htg26: *
76: gb_htg27: *
77: gb_htg28: *
78: gb_htg29: *
79: gb_htg30: *
80: gb_htg31: *
81: gb_v11: *
82: gb_v12: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	13	100.0	487	9 AF179732	AF179732 Pan trogl
2	13	100.0	904	4 AF027222	AF027222 Cephus g
3	13	100.0	904	4 AF027231	AF027231 Cephus g
4	13	100.0	906	4 AF027225	AF027225 Cephus g
5	13	100.0	910	4 AF027233	AF027233 Cephus g
6	13	100.0	911	4 AF027226	AF027226 Cephus g
7	13	100.0	911	4 AF027227	AF027227 Cephus g
8	13	100.0	911	4 AF027228	AF027228 Cephus g
9	13	100.0	915	4 AF027230	AF027230 Cephus g
10	13	100.0	939	4 AF027231	AF027231 Cephus g
11	13	100.0	1242	11 AF022953	AF022953 Homo sapi
12	13	100.0	1290	39 AF169225	AF169225 Homo sapi
13	13	100.0	1290	39 AF203386	AF203386 Homo sapi
14	13	100.0	1340	33 CEPHOCAR	CEPHOCAR
15	13	100.0	1340	33 CEPHOCAR	CEPHOCAR
16	13	100.0	1959	12 RATYADBC	RATYADBC
17	13	100.0	1970	9 HSBAR	HSBAR
18	13	100.0	2165	33 AAU86080	AAU86080 Aedes aegypt
19	13	100.0	2339	10 AF047430	AF047430 Homo sapi
C 20	13	100.0	2426	2 AF047430	AF047430 Homo sapi
21	13	100.0	3458	10 HUMADBRRA	HUMADBRRA
22	13	100.0	3780	12 RNO35448	RNO35448 Rattus norv
23	13	100.0	4197	12 RSB2AR	RSB2AR
24	13	100.0	4343	12 MMTCAN	MMTCAN

25 13 100.0 4446 1 PSP710POL 254173 Pyrococcus
26 13 100.0 4446 5 A68743 A68743 Sequence 1
27 13 100.0 5506 9 HSPPOXG X99450 H. sapiens p
28 13 100.0 6330 11 AF037062 AF037062 Homo sapi
29 13 100.0 13558 8 AF093161 U93161 Arabidopsis
30 13 100.0 20891 33 CERF01G4 Z68341 Caenorhabdi
31 13 100.0 36815 33 CERF35H10 U40934 Caenorhabdi
32 13 100.0 37923 33 CELK07H8 AF047659 Caenorhab
33 13 100.0 40573 11 HSN15422 U15422 Human Prota
34 13 100.0 49474 76 AC055838 AC055838 Homo sapi
35 13 100.0 59498 33 AC004252 AC004252 Drosophill
36 13 100.0 67476 43 AC021969 AC021969 Homo sapi
37 13 100.0 71244 43 AC0021326 AC0021326 Homo sapi
38 13 100.0 71504 33 AC002473 AC002473 Arabidops
39 13 100.0 76526 72 AC034106 AC034106 Arabidops
40 13 100.0 82879 51 AC022656 AC022656 Homo sapi
41 13 100.0 84034 12 AC006943 AC006943 Mus muscu
42 13 100.0 88933 42 AC017870 AC017870 Drosophill
43 13 100.0 93214 42 AC013107 AC013107 Drosophill
44 13 100.0 94786 49 AC002333 AC002333 Arabidops
45 13 100.0 99033 11 AC005704 AC005704 Homo sapi

ALIGNMENTS

RESULT 1
AF179732 487 bp DNA PRI 09-MAR-2000
LOCUS Pan troglodytes olfactory receptor (PTR207) gene, partial cds.
DEFINITION AF179732
VERSION AF179732.1 GI:7211478

SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
AUTHORS Rouquier,S., Blancher,A. and Giorgi,D.
TITLE The olfactory receptor gene repertoire in primates and mouse:
evidence for reduction of the functional fraction in primates
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (6), 2870-2874 (2000)
MEDLINE 20183981
REFERENCE 2 (bases 1 to 487)
AUTHORS Giorgi,D.G. and Rouquier,S.P.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-1999) Institut de Genetique Humaine, CNRS, UPR
1142, rue de la Cardonille, Montpellier Cedex 5 34396, France

FEATURES
Source Location/Qualifiers
1..487

gene /organism="Pan troglodytes"
/db_xref="taxon:9598"
CDS <1..>487
<1..>487
/gene="PTR207"
/codon_start=2
/product="olfactory receptor"
/protein_id="AA040327.1"
/db_xref="GI:7211479"
/translation="VAVCNPLVTVMYORLCSLVATSCGRVCSLTLYELLEIS
FRGNININNVCEHAAIVAVSCDPVQSEITLVSAFNEISLVITLSTAFIPITV
MKTASIGCRKKAFETCASHLTATITFHFGLTFLYCVNKSRLWAKVASVEYTVVTP
ML"

BASE COUNT 99 a 122 c 103 g 163 t

ORIGIN

Query Match 100.0%; Score 13; DB 9; Length 487;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0y 1 cccaatggaagcc 13

Db 328 CCAATGGAAGCC 316

RESULT 2

LOCUS AF027232 904 bp DNA VRT 08-JUL-1998
DEFINITION Cepphus grylle arcticus isolate CGA-SD-0001 mitochondrial control
region, partial sequence.
ACCESSION AF027232
VERSION AF027232.1 GI:3264880

KEYWORDS Cepphus grylle arcticus.
ORGANISM Mitochondrion Cepphus grylle arcticus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Charadriiformes; Alcidae; Cepphus.

REFERENCE
AUTHORS Kidd,M.G. and Friesen,V.L.
TITLE Sequence variation in the guillemot (Alcidae: Cepphus)
mitochondrial control region and its nuclear homolog
JOURNAL Mol. Biol. Evol. 15 (1), 61-70 (1998)
MEDLINE 98152305

REFERENCE 2 (bases 1 to 904)
AUTHORS Kidd,M.G. and Friesen,V.L.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1997) Biology, Queen's University, Kingston, ON
K7L 3N6, Canada

FEATURES
Source Location/Qualifiers
1..904

misc.feature
BASE COUNT 229 a 229 c 142 g 258 t 46 others
ORIGIN
/organism="Cepphus grylle arcticus"
/organelle="mitochondrion"
/isolate="CGA-SD-0001"
/sub_species="arcticus"
/db_xref="taxon:78224"
/note="common name: black guillemot"
<1..>904
/note="mitochondrial control region"

Query Match 100.0%; Score 13; DB 4; Length 904;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0y 1 cccaatggaagcc 13
Db 15 CCAATGGAAGCC 27

RESULT 3

LOCUS AF027251 904 bp DNA VRT 08-JUL-1998
DEFINITION Cepphus grylle grylle isolate CCG-SF-0001 mitochondrial control
region, partial sequence.
ACCESSION AF027251
VERSION AF027251.1 GI:3264899

KEYWORDS Cepphus grylle grylle.
ORGANISM Mitochondrion Cepphus grylle grylle
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Charadriiformes; Alcidae; Cepphus.
REFERENCE 1 (bases 1 to 904)
AUTHORS Kidd,M.G. and Friesen,V.L.
TITLE Sequence variation in the guillemot (Alcidae: Cepphus)
mitochondrial control region and its nuclear homolog
JOURNAL Mol. Biol. Evol. 15 (1), 61-70 (1998)
MEDLINE 98152305

REFERENCE 2 (bases 1 to 904)
AUTHORS Kidd,M.G. and Friesen,V.L.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-1997) Biology, Queen's University, Kingston, ON
K7L 3N6, Canada

FEATURES
source
1. 904
/organism="Cepphus grylle grylle"
/organelle="mitochondrion"
/isolate="CG-SD-0001"
/sub_species="grylle"
/db_xref="taxon:78225"
/note="common name: black guillemot"
1..>904
/note="mitochondrial control region"
BASE COUNT 228 a 238 c 143 g 277 t 18 others
ORIGIN

Query Match 100.0%; Score 13; DB 4; Length 904;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagc 13
|||||
DB 15 CCCAATGGAAGC 27

RESULT 4
AF027225 906 bp DNA VRT 08-JUL-1998
LOCUS
DEFINITION Cepphus grylle ultimus isolate CGU-DS-2189 mitochondrial control region, partial sequence.
ACCESSION AF027225
VERSION AF027225.1 GI:3264873
KEYWORDS
SOURCE
ORGANISM
Cepphus grylle ultimus.
Mitochondrion Cepphus grylle ultimus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Charadriiformes; Alcidae; Cepphus.
REFERENCE
1 (bases 1 to 906)
Kidd,M.G. and Friesen,V.L.
Sequence variation in the guillemot (Alcidae: Cepphus)
mitochondrial control region and its nuclear homolog
Mol. Biol. Evol. 15 (1), 61-70 (1998)
JOURNAL
MEDLINE
98152305
REFERENCE
2 (bases 1 to 906)
Kidd,M.G. and Friesen,V.L.
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (29-SEP-1997) Biology, Queen's University, Kingston, ON
K7L 3N6, Canada
FEATURES
source
Location/Qualifiers
1. 906
/organism="Cepphus grylle ultimus"
/organelle="mitochondrion"
/isolate="CGU-DS-2189"
/sub_species="ultimus"
/db_xref="taxon:78228"
/note="common name: black guillemot"
1..>906
/note="mitochondrial control region"
BASE COUNT 234 a 242 c 137 g 278 t 15 others
ORIGIN

Query Match 100.0%; Score 13; DB 4; Length 906;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagc 13
|||||
DB 10 CCCAATGGAAGC 22

RESULT 5
AF027233

LOCUS AF027233 910 bp DNA VRT 08-JUL-1998
DEFINITION Cepphus grylle arcticus isolate CGA-SD-0002 mitochondrial control region, partial sequence.
ACCESSION AF027233
VERSION AF027233.1 GI:3264881
KEYWORDS
SOURCE
ORGANISM
Cepphus grylle arcticus.
Mitochondrion Cepphus grylle arcticus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Charadriiformes; Alcidae; Cepphus.
REFERENCE
1 (bases 1 to 910)
Kidd,M.G. and Friesen,V.L.
Sequence variation in the guillemot (Alcidae: Cepphus)
mitochondrial control region and its nuclear homolog
Mol. Biol. Evol. 15 (1), 61-70 (1998)
JOURNAL
MEDLINE
98152305
REFERENCE
2 (bases 1 to 910)
Kidd,M.G. and Friesen,V.L.
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (29-SEP-1997) Biology, Queen's University, Kingston, ON
K7L 3N6, Canada
FEATURES
source
Location/Qualifiers
1. 910
/organism="Cepphus grylle arcticus"
/organelle="mitochondrion"
/isolate="CGA-SD-0002"
/sub_species="arcticus"
/db_xref="taxon:78224"
/note="common name: black guillemot"
1..>910
/note="mitochondrial control region"
BASE COUNT 232 a 236 c 141 g 271 t 30 others
ORIGIN

Query Match 100.0%; Score 13; DB 4; Length 910;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagc 13
|||||
DB 15 CCCAATGGAAGC 27

RESULT 6
AF027226 911 bp DNA VRT 08-JUL-1998
LOCUS
DEFINITION Cepphus grylle ultimus isolate CGU-DS-2190 mitochondrial control region, partial sequence.
ACCESSION AF027226
VERSION AF027226.1 GI:3264874
KEYWORDS
SOURCE
ORGANISM
Cepphus grylle ultimus.
Mitochondrion Cepphus grylle ultimus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Charadriiformes; Alcidae; Cepphus.
REFERENCE
1 (bases 1 to 911)
Kidd,M.G. and Friesen,V.L.
Sequence variation in the guillemot (Alcidae: Cepphus)
mitochondrial control region and its nuclear homolog
Mol. Biol. Evol. 15 (1), 61-70 (1998)
JOURNAL
MEDLINE
98152305
REFERENCE
2 (bases 1 to 911)
Kidd,M.G. and Friesen,V.L.
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (29-SEP-1997) Biology, Queen's University, Kingston, ON
K7L 3N6, Canada
FEATURES
source
Location/Qualifiers
1. 911
/organism="Cepphus grylle ultimus"
/organelle="mitochondrion"
/isolate="CGU-DS-2190"
/sub_species="ultimus"

```

/db_xref="taxon:78228"
/note="common name: black guillemot"
misc_feature
<1..>911
/note="mitochondrial control region"
misc_feature
823..857
/note="conserved sequence block-1"
BASE COUNT      238 a      239 c      138 g      270 t      26 others
ORIGIN

Query Match      100.0%; Score 13; DB 4; Length 911;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cccaatggaagc 13
        |||
        15 CCCAATGGAAGC 27

RESULT      7
AF027227      911 bp      DNA      VRT      08-JUL-1998
LOCUS      Cephus grylle islandicus isolate CGI-FI-002A mitochondrial control
DEFINITION      region, partial sequence.
ACCESSION      AF027227.1 GI:3264875
VERSION
KEYWORDS
SOURCE
ORGANISM      Cephus grylle islandicus.
               Mitochondrion Cephus grylle islandicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
               Aves; Neognathae; Charadriiformes; Alcidae; Cephus.
REFERENCE
AUTHORS      Kidd, M.G. and Friesen, V.L.
TITLE      Sequence variation in the guillemot (Alcidae: Cephus)
               mitochondrial control region and its nuclear homolog
MOL. BIOL. EVOL. 15 (1), 61-70 (1998)
JOURNAL
MEDLINE
REFERENCE
AUTHORS      Kidd, M.G. and Friesen, V.L.
TITLE      Direct Submission
SUBMITTED (29-SEP-1997) Biology, Queen's University, Kingston, ON
JOURNAL
FEATURES
SOURCE      location/Qualifiers
               1..911
               /organism="Cephus grylle islandicus"
               /organella="mitochondrion"
               /isolate="CGI-FI-002A"
               /sub_species="islandicus"
               /db_xref="taxon:78226"
               /note="common name: black guillemot"
               <1..>911
               /note="mitochondrial control region"
BASE COUNT      230 a      237 c      137 g      266 t      41 others
ORIGIN

Query Match      100.0%; Score 13; DB 4; Length 911;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cccaatggaagc 13
        |||
        15 CCCAATGGAAGC 27

RESULT      8
AF027228      911 bp      DNA      VRT      08-JUL-1998
LOCUS      Cephus grylle islandicus isolate CGI-FI-002B mitochondrial control
DEFINITION      region, partial sequence.
ACCESSION      AF027228
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

```

```

VERSION      AF027228.1 GI:3264876
KEYWORDS
SOURCE
ORGANISM      Cephus grylle islandicus.
               Mitochondrion Cephus grylle islandicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
               Aves; Neognathae; Charadriiformes; Alcidae; Cephus.
REFERENCE
AUTHORS      Kidd, M.G. and Friesen, V.L.
TITLE      Sequence variation in the guillemot (Alcidae: Cephus)
               mitochondrial control region and its nuclear homolog
MOL. BIOL. EVOL. 15 (1), 61-70 (1998)
JOURNAL
MEDLINE
REFERENCE
AUTHORS      Kidd, M.G. and Friesen, V.L.
TITLE      Direct Submission
SUBMITTED (29-SEP-1997) Biology, Queen's University, Kingston, ON
JOURNAL
FEATURES
SOURCE      location/Qualifiers
               1..911
               /organism="Cephus grylle islandicus"
               /organella="mitochondrion"
               /isolate="CGI-FI-002B"
               /sub_species="islandicus"
               /db_xref="taxon:78226"
               /note="common name: black guillemot"
               <1..>911
               /note="mitochondrial control region"
BASE COUNT      227 a      228 c      138 g      268 t      50 others
ORIGIN

Query Match      100.0%; Score 13; DB 4; Length 911;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cccaatggaagc 13
        |||
        15 CCCAATGGAAGC 27

RESULT      9
AF027230      915 bp      DNA      VRT      08-JUL-1998
LOCUS      Cephus grylle mandtii isolate CGM-SV-0001 mitochondrial control
DEFINITION      region, partial sequence.
ACCESSION      AF027230.1 GI:3264878
VERSION
KEYWORDS
SOURCE
ORGANISM      Cephus grylle mandtii.
               Mitochondrion Cephus grylle mandtii
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
               Aves; Neognathae; Charadriiformes; Alcidae; Cephus.
REFERENCE
AUTHORS      Kidd, M.G. and Friesen, V.L.
TITLE      Sequence variation in the guillemot (Alcidae: Cephus)
               mitochondrial control region and its nuclear homolog
MOL. BIOL. EVOL. 15 (1), 61-70 (1998)
JOURNAL
MEDLINE
REFERENCE
AUTHORS      Kidd, M.G. and Friesen, V.L.
TITLE      Direct Submission
SUBMITTED (29-SEP-1997) Biology, Queen's University, Kingston, ON
JOURNAL
FEATURES
SOURCE      location/Qualifiers
               1..915
               /organism="Cephus grylle mandtii"
               /organella="mitochondrion"
               /isolate="CGM-SV-0001"
               /sub_species="mandtii"
               /db_xref="taxon:78227"
               /note="common name: black guillemot"

```

```

misc_feature <1..>915
              /note="mitochondrial control region"
BASE COUNT   235 a      243 c      136 g      264 t      37 others
ORIGIN
Query Match          100.0%; Score 13; DB 4; Length 915;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
    |||
Db 12 CCCAATGGAAGCC 24

RESULT 10
LOCUS AF027231 939 bp DNA VRT 08-JUL-1998
DEFINITION Cephus grylle mandill isolate CGM-SV-0002 mitochondrial control
ACCESSION AF027231 GI:3264879
VERSION AF027231.1 GI:3264879
KEYWORDS
SOURCE Cephus grylle mandill.
          Mitochondrion Cephus grylle mandill
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauaria;
          Aves; Neognathae; Charadriiformes; Alcidae; Cephus.
REFERENCE 1 (bases 1 to 939)
AUTHORS Kidd,M.G. and Friesen,V.L.
TITLE Sequence variation in the guillemot (Alcidae: Cephus)
          mitochondrial control region and its nuclear homolog
JOURNAL MOL. BIOL. EVOL. 15 (1), 61-70 (1998)
MEDLINE 98152305
REFERENCE 2 (bases 1 to 939)
AUTHORS Kidd,M.G. and Friesen,V.L.
TITLE Direct Submision
JOURNAL Submitted (29-SEP-1997) Biology, Queen's University, Kingston, ON
          K7L 3N6, Canada
FEATURES
    source
        1..939
        /organism="Cephus grylle mandill"
        /organella="mitochondrion"
        /isolate="CGM-SV-0002"
        /sub_species="mandill"
        /db_xref="taxon:78227"
        /note="common name: black guillemot"
        <1..>939
        /note="mitochondrial control region"
BASE COUNT   260 a      250 c      144 g      276 t      9 others
ORIGIN
Query Match          100.0%; Score 13; DB 4; Length 939;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
    |||
Db 15 CCCAATGGAAGCC 27

RESULT 11
LOCUS AF022953 1242 bp DNA PRI 30-OCT-1997
DEFINITION Homo sapiens beta2-adrenergic receptor (ADRB2) gene, complete cds.
ACCESSION AF022953
VERSION AF022953.1 GI:2570526
KEYWORDS
SOURCE human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1242)

```

```

AUTHORS Reibhaus,E., Innis,M., Macintyre,N. and Liggett,S.B.
TITLE Mutations in the gene encoding for the beta 2-adrenergic receptor
          in normal and asthmatic subjects
JOURNAL Am. J. Respir. Cell Mol. Biol. 8 (3), 334-339 (1993)
MEDLINE 93192047
REFERENCE 2 (bases 1 to 1242)
AUTHORS Green,S.A., Cole,G., Jacinto,M., Innis,M. and Liggett,S.B.
TITLE A polymorphism of the human beta 2-adrenergic receptor within the
          fourth transmembrane domain alters ligand binding and functional
          properties of the receptor
JOURNAL J. Biol. Chem. 268 (31), 23116-23121 (1993)
MEDLINE 94043092
REFERENCE 3 (bases 1 to 1242)
AUTHORS Green,S.A., Turki,J., Innis,M. and Liggett,S.B.
TITLE Amino-terminal polymorphisms of the human beta 2-adrenergic
          receptor impart distinct agonist-promoted regulatory properties
JOURNAL Biochemistry 33 (32), 9414-9419 (1994)
MEDLINE 94347707
REMARK Erratum:[published erratum appears in Biochemistry 1994 Nov
          29;33(47):14368]]
REFERENCE 4 (bases 1 to 1242)
AUTHORS Liggett,S.B. and Green,S.A.
TITLE Direct Submision
JOURNAL Submitted (04-SEP-1997) Medicine, Univ of Cincinnati, 231 Bethesda
          Ave ML670564, Cincinnati, OH 45267-0564, USA
FEATURES
    source
        1..1242
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        1..1242
        /gene="ADRB2"
        1..1242
        /gene="ADRB2"
        /codon_start=1
        /product="beta2-adrenergic receptor"
        /protein_id="AAB82148.1"
        /db_xref="GI:2570527"
        /translation="MGCPGNGSAFLIAPGSHAPDHDTQORDEWVYVNGTIVMSLI
          LAIVGNVIVITAIKFERLQVTVNFTSLKADLVKGLAVPPGAAHIILKMKTFG
          NFWCEWISIDVLCVTASTETLCLVAVDVFATITSPFKYOSLTNNKAVIILMWIV
          SGLTSLPLQIMHMYRATHQEAINCYANECDFNQAVAIASISIVSPVPLVIVFV
          YSRVPEAKRQLOKIDKSEGRHVQMLSGVEDGRGGRGRLRSSKFKLSEKALKLTIG
          IIMGFTLCMLPPEIVIVIVIODNIRREYVILMWIGIVNSGFNPPLYCHSPDR
          APOELLCARSSSLKAYGNGSYNGNIGDSGYHVEDEKKNKLCELDLPETDPVCHGS
          TVPSDNIDSGKNCSTNDSL.
          46
        /gene="ADRB2"
        /note="Arg16 to Gly polymorphism"
        /replace="a"
BASE COUNT   275 a      331 c      326 g      310 t
ORIGIN
Query Match          100.0%; Score 13; DB 11; Length 1242;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
    |||
Db 40 CCCAATGGAAGCC 52

RESULT 12
LOCUS AF169225 1290 bp DNA PRI 10-AUG-1999
DEFINITION Homo sapiens beta-2-adrenergic receptor gene, complete cds.
ACCESSION AF169225
VERSION AF169225.1 GI:5714687
KEYWORDS
SOURCE human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
          Eutheria; Primates; Catarrhini; Homidae; Homo.

```

```

REFERENCE      1 (bases 1 to 1290)
AUTHORS        Rupert,J.R. and Hochachka,P.W.
TITLE          Beta-2-adrenergic receptor allele frequencies in two native
               American populations
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 1290)
AUTHORS        Rupert,J.R. and Hochachka,P.W.
TITLE          Direct Submission
JOURNAL        Submitted (14-JUL-1999) Zoology, University of British Columbia,
               6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
FEATURES       source
               1..1290
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="5"
               /map="5q31-q33"
               /cell_type="lymphocyte"
               /tissue_type="blood"
               /note="Isolated from a Quechua speaking native American
               heterozygous for a known C/T mutation"
               17..1258
               /codon_start=1
               /product="beta-2-adrenergic receptor"
               /protein_id="A048036.1"
               /db_xref="GI:5714688"
               /translation="MGQPGNGSAFLAPNGSHAPDHYTOORDEWVYVGMGIVMSLIY
               LAIVGNAVLYTALAKFERLQVTNYNPTISLACADLVWGLAVPEGAHILMKMTFG
               NFWCEFTSIDVLCVASTIEYLAVDRFAITSPKQSLTKNKARVITLMWTV
               SGLTSFLPIOMHMYRATHQEAINCYNATCCDFTNOAYALASIVSFYVLIVWFV
               YSRVFOEAROLOKIDKSGRPHVONLSQVEDGRTGGLRRSKFKLEKALKTLG
               IIMGFTLCWLPFTYNIYVHVDNLIRKEYILLNMGVYNSGFNPLICRSPDFRI
               AFDELICLRSSSLKAYNGSSNGTNGSGYHVDEKRNKLLCEDLPGETDFVGHOG
               TVPSDNIDSGRNCSTNDSL"
               491
variation      /replace="C/T"
BASE COUNT    287 a      349 c      331 g      322 t      1 others
ORIGIN
Query Match   100.0%; Score 13; DB 39; Length 1290;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cccaatggaacc 13
    |||||
Db 56 CCCAATGGAACC 68
RESULT 13
AF203386 1290 bp DNA PRI 28-DEC-1999
LOCUS     Homo sapiens beta-2 adrenergic receptor (ADRB2) gene, complete cds.
DEFINITION
ACCESSION AF203386
VERSION    AF203386.1 GI:6636495
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Euthera; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 1290)
AUTHORS    Rupert,J.L. and Hochachka,P.W.
TITLE      Beta-2 adrenergic receptor allele frequencies in two Native
            American populations
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1290)
AUTHORS    Rupert,J.L. and Hochachka,P.W.
TITLE      Direct Submission
JOURNAL    Submitted (09-NOV-1999) Zoology, University of British Columbia,
            6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
FEATURES    source
            1..1290
            /organism="Homo sapiens"
            /db_xref="taxon:9606"

```

```

               /chromosome="5"
               /map="5q32-q34"
               /sex="female"
               /note="Isolated from a Quechuan-speaking Native American"
               26..1267
               /gene="ADRB2"
               26..1267
               /gene="ADRB2"
               /note="catecholamine receptor"
               /codon_start=1
               /product="beta-2 adrenergic receptor"
               /protein_id="AAF20199.1"
               /db_xref="GI:6636496"
               /translation="MGQPGNGSAFLAPNGSHAPDHYTOORDEWVYVGMGIVMSLIY
               LAIVGNAVLYTALAKFERLQVTNYNPTISLACADLVWGLAVPEGAHILMKMTFG
               NFWCEFTSIDVLCVASTIEYLAVDRFAITSPKQSLTKNKARVITLMWTV
               SGLTSFLPIOMHMYRATHQEAINCYNATCCDFTNOAYALASIVSFYVLIVWFV
               YSRVFOEAROLOKIDKSGRPHVONLSQVEDGRTGGLRRSKFKLEKALKTLG
               IIMGFTLCWLPFTYNIYVHVDNLIRKEYILLNMGVYNSGFNPLICRSPDFRI
               AFDELICLRSSSLKAYNGSSNGTNGSGYHVDEKRNKLLCEDLPGETDFVGHOG
               TVPSDNIDSGRNCSTNDSL"
BASE COUNT    288 a      345 c      333 g      324 t
ORIGIN
Query Match   100.0%; Score 13; DB 39; Length 1290;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cccaatggaacc 13
    |||||
Db 65 CCCAATGGAACC 77
RESULT 14
CEPHOCAR 1340 bp mRNA INV 03-AUG-1995
LOCUS     C.elegans mRNA for phosphate carrier protein.
DEFINITION
ACCESSION X76113
VERSION    X76113.1 GI:472905
KEYWORDS   phosphate carrier.
SOURCE     Caenorhabditis elegans.
ORGANISM   Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Secernentea; Rhabdita; Rhabditidae;
            Rhabditina; Rhabdilloidea; Rhabdilitidae; Pelodierinae; Caenorhabditis.
REFERENCE  1 (bases 1 to 1340)
AUTHORS    Walker,J.E.
TITLE      Direct Submission
JOURNAL    Submitted (10-NOV-1993) J.E. Walker, MRC Lab of Molecular Biology,
            Hills Road, Cambridge CB2 2QH, UK
REFERENCE  2 (bases 1 to 1340)
AUTHORS    Runswick,M.J., Philippides,A., Lauria,G. and Walker,J.E.
TITLE      Extension of the mitochondrial transport superfamily: sequences of
            five members from the nematode worm Caenorhabditis elegans
JOURNAL    Unpublished
REFERENCE  3 (bases 1 to 1340)
AUTHORS    Runswick,M.J., Philippides,A., Lauria,G. and Walker,J.E.
TITLE      Extension of the mitochondrial transport superfamily: sequences
            of five members from the nematode worm, Caenorhabditis elegans
JOURNAL    DNA Seq. 4 (5), 281-291 (1994)
FEATURES    source
            1..1340
            /organism="Caenorhabditis elegans"
            /db_xref="taxon:6239"
            /genome="germline"
            /clone_lib="lambda"
            5..1027
            /codon_start=1
            /product="phosphate carrier protein"
            /protein_id="CAA53719.1"
            /db_xref="GI:472906"
            /db_xref="SWISS-PROT:P40614"

```

/translation="MSVFSQLAESSKONPFSLPVRSNCASAVSAPGVFEGSGKYA
YCALGVLSCGITHFAIVPLDLVKCRIVNPEKRYGIATGPTTIAEBCARALVGM
PTLIGYSAGLGFGEYEFKKNVYADMDGEENALYRTSLYLAASAEFPADLLAP
MEATYRIOTSPSCAPPTLRGCAPMITYKAGLPGFKGLEPLMKROIPTYMKPFACREK
TYEALYVYVPRAPRACSKAEQLVTFVAGTLAGFCALVSPADIVASKLNDOSAT
AGGILKKLGFAGVWKGIVPRITMIGTLALQWFIYDSVAVNALNLPPEPEMPASLKA
KLAADQ"

translt_peptide 8.85
mat_peptide 86.1024
polya_site /product="phosphate carrier protein"
1312.1340
BASE COUNT 347 a 359 c 268 g 366 t
ORIGIN

Query Match 100.0%; Score 13; DB 33; Length 1340;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagcc 13
|||||
Db 481 CCCAATGGAAGCC 493

RESULT 15
RATADBC RATADBC 1959 bp mRNA ROD 27-APR-1993
DEFINITION Rat beta-adrenergic receptor mRNA, complete cds.
ACCESSION J03024.1 GI:202710
VERSION J03024.1
KEYWORDS beta-adrenergic receptor.
SOURCE Rat heart, cDNA to mRNA, (library of Clontech), clone
lambda-RHB-DAR.

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 1959)
Gocayne, J., Robinson, D.A., Fitzgerald, M.G., Chung, F.-Z.,
Kerlavage, A.R., Lentes, K.-U., Lai, J., Wang, C.-D., Fraser, C.M. and
Venter, J.C.
Primary structure of rat cardiac beta-adrenergic and muscarinic
cholinergic receptors obtained by automated DNA sequence analysis:
Further evidence for a multigene family
Proc. Natl. Acad. Sci. U.S.A. 84, 8296-8300 (1987)
88068581

COMMENT Draft entry and computer-readable sequence for [1] kindly provided
by A.Kerlavage, 20-OCT-1987.

FEATURES
source
1. 1959
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
<1..1959
/note="ADB mRNA"
102..1358
/note="beta-adrenergic receptor"
/codon_start=1
/protein_id="AAA40675.1"
/db_xref="GI:202711"
/translation="MEPHGNDSDFLAPNGSAPGHDTIOERDEAMVYGMALIMSIVT
LAIVGNAVITAIKPERLQVTNYFTISLACADLVMLGAVPFGASHILKMNFG
NFMCEFWTSIDVLCVTASIECTVAVDRYVAITSPFKQSLLTKNKARAVVLTAMV
SGLTSEFLPIOMHWYRATHKQADIDYAKETCODEFTNOAVAIASSIVSFVPLVWVAV
YSRROVAKROLOKIDKSGRPHQNLQSOVEDODSGSGCLRSSKFKLKEHAKLTG
IIMGTFTLCMLPFIYVNYVIRANDIKREVITILNMLGYNSARNPILYCRSPDRI
AFQELCLARRSSSKTYGNGYSSNGRTDYGESAYOLGQEKELLCEAPGMEGF
VNCQGTPLSLIDSGRNCNTNDSPL"

BASE COUNT 466 a 476 c 492 g 525 t
ORIGIN 1 bp upstream of HindIII site.

Query Match 100.0%; Score 13; DB 12; Length 1959;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagcc 13
|||||
Db 141 CCCAATGGAAGCC 153

Search completed: September 12, 2000, 23:03:24
Job time: 3996 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:07:23 ; Search time 148.16 Seconds
(without alignments)
21.953 Million cell updates/sec

Title: US-09-542-718-5
Perfect score: 13
Sequence: 1 cccaatggaagcc 13

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 623170

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	13	100.0	4446 1	DNA encoding a DNA
2	12	92.3	69 1	STAPHYLOCOCCUS aur
3	12	92.3	319 1	Nucleic acid G30.
4	12	92.3	501 1	Enterococcus faeca
5	12	92.3	1041 1	Gene encoding Elk-
6	12	92.3	1041 1	Human cytokine elk
7	12	92.3	1041 1	Human elk ligand c
8	12	92.3	1041 1	Human elk-L coding
9	12	92.3	1041 1	Human elk-L encodi
10	12	92.3	1086 1	Cellulose synthase
11	12	92.3	1516 1	H. pylori GHP0 701
12	12	92.3	1575 1	Sequence of the as
13	12	92.3	1575 1	ADPGP small subun
14	12	92.3	1575 1	ADPGP small subun
15	12	92.3	1836 1	Brevibacterium fla
16	12	92.3	2015 1	DNA encoding Marek
17	12	92.3	2170 1	Yeast TFI1 gene. I
18	12	92.3	2625 1	S.cerevisiae TFI1
19	12	92.3	2625 1	S.cerevisiae TFI1
20	12	92.3	2625 1	Human canalicular
21	12	92.3	4868 1	Human canalicular
22	12	92.3	4993 1	HincII fragment of
23	12	92.3	4993 1	HincII fragment of
24	12	92.3	5386 1	Human canalicular
25	12	92.3	6923 1	Polynucleotide seq
26	12	92.3	21126 1	TL-DNA region of A
27	11.4	87.7	20 1	Actinobacillus act
28	11.4	87.7	20 1	Actinobacillus act
29	11.4	87.7	20 1	Actinobacillus act
30	11.4	87.7	20 1	Actinobacillus act
31	11.4	87.7	20 1	Actinobacillus act
32	11.4	87.7	22 1	Actinobacillus act
33	11.4	87.7	24 1	Actinobacillus act

ALIGNMENTS

34	11.4	87.7	26 1	077909	Actinobacillus act
35	11.4	87.7	28 1	077910	Actinobacillus act
36	11.4	87.7	240 1	T20105	Human gene signatu
37	11.4	87.7	270 1	T69336	Murine metastatic
38	11.4	87.7	307 1	Q39662	Expressed Sequence
39	11.4	87.7	309 1	Q59074	Human brain expres
40	11.4	87.7	309 1	T69333	Murine metastatic
41	11.4	87.7	347 1	Q60393	Human brain expres
42	11.4	87.7	400 1	V76412	Staphylococcus aur
43	11.4	87.7	422 1	X41467	Human secreted pro
44	11.4	87.7	445 1	T19058	Human gene signatu
45	11.4	87.7	473 1	T69342	Murine metastatic

RESULT 1

ID	V36034	Standard: DNA; 4446 BP.
AC	V36034;	
DT	02-SEP-1998 (first entry)	
DE	DNA encoding a DNA polymerase of Pyrococcus sp. GE23.	
KM	Heat stable DNA polymerase; archaeobacteria; secondary structure; ds.	
OS	Pyrococcus sp.	
FH	Key	
FT	CDS	
FT	Location/Qualifiers	
FT	/*lag- a	
FT	/product- DNA polymerase	
PN	MO9801567-A2.	
PD	15-JUN-1998.	
PF	10-JUL-1997; F01259.	
PR	10-JUL-1996; FR-008631.	
PA	(APPL-) APPLIGENE-ONCOR SA.	
PI	Camdon MA, Querrelon J;	
DR	WPI: 98-101062/09.	
DR	P-SDS: W60719.	
PT	Heat stable DNA polymerase from Pyrococcus species - and related	
PT	DNA, vectors and transformed cells, useful in nucleic acid	
PT	amplification reactions performed at temperatures high enough to	
PT	denature secondary structures	
PS	Claim 9; Pages 37-41; 80pp; French.	
CC	The present sequence encodes a heat stable DNA polymerase from	
CC	archaeobacteria of the genus Pyrococcus sp. GE23. The encoded polymerase	
CC	has an extremely high heat stability and can catalyze reactions at	
CC	70-90 degrees Celsius. High molecular weight DNA was isolated from the	
CC	bacteria, and screened with probes based on the polymerase genes of	
CC	P. furiosus and Thermococcus litoralis to isolate the present sequence.	
CC	Host cells transformed with DNA encoding the present polymerase are	
CC	used to produce recombinant DNA which is useful for nucleic acid	
CC	amplification, including direct, reversed polymerase chain reaction (PCR)	
CC	on double-stranded matrices and PCR on matrices having secondary	
CC	structures that block the process at conventional elongation	
CC	temperatures.	
SO	Sequence 4446 BP; 1358 A; 884 C; 1115 G; 1089 T;	

Query Match 100.0%; Score 13; DB 1; Length 4446;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2527 CCCAATGGAAGCC 2539

RESULT 2

ID	V76506/C	Standard: DNA; 69 BP.
AC	V76506;	
DT	16-MAR-1999 (first entry)	
DE	Staphylococcus aureus contig SEQ ID #2195.	
KM	Computer readable medium; vaccine; S.aureus infection; immunodetection;	

KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 OS Staphylococcus aureus.
 PN Ep-786519-A2.
 PD 30-JUL-1997.
 PF 07-JAN-1997; 100117.
 PR 05-JAN-1996; US-009861.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA,
 PI Rosen CA,
 PI WPI: 97-374922/35.
 DR Polynucleotide(s) and proteins derived from Staphylococcus aureus
 PT stored on computer readable medium and used in the production of
 PT anti-S.aureus vaccines
 PS Claim 1: Page 2172: 3271bp; English.
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access
 CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 CC the S.aureus DNA sequences allows putative functions to be assigned so
 CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S.aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S.aureus DNA sequences contained on the
 CC computer readable medium.
 SQ Sequence 69 BP; 19 A; 10 C; 18 G; 22 T;

Query Match 92.3%; Score 12; DB 1; Length 69;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagc 12
 DB 30 CCCAATGGAAGC 19

RESULT 3
 V60051/C
 ID V60051 standard; cDNA to mRNA; 319 BP.
 AC V60051;
 DT 25-NOV-1998 (first entry)
 DE Nucleic acid 630.
 KW Human homologue; yeast RAD50; Drosophila Septin-2; Acyl-CoA synthetase;
 KW immunomodulatory activity; identification; activated T-cell; ds.
 OS Homo sapiens.
 PN W09838306-A1.
 PD 03-SEP-1998.
 PR 27-FEB-1997; U03159.
 PR 27-FEB-1997; WO-U03159.
 PA (GENE-) GENELABS TECHNOLOGIES INC.
 PI Dolganov G;
 PI WPI: 98-481207/41.
 DR Novel human immunomodulatory poly(peptide)s - have homology to the
 PT yeast RAD50 or Drosophila Septin-2 proteins
 PS Disclosure: Page 119; 153bp; English.
 CC The present sequence represents a nucleic acid sequence identified in
 CC the course of the invention, and expressed in lung, brain, kidney, heart,
 CC muscle, liver, placenta, small intestine and activated T-cells. The
 CC specification also describes sequences encoding human homologues of the
 CC yeast RAD50, the Drosophila Septin-2 and Acyl-CoA synthetase. The
 CC proteins have immunomodulatory activity. The nucleic acids and proteins
 CC can be used to identify activated T-cells in a sample population. They
 CC can also be used to isolate and identify sequences encoding other

CC proteins or other compounds having immunomodulatory activity.
 SQ Sequence 319 BP; 96 A; 56 C; 74 G; 93 T;

Query Match 92.3%; Score 12; DB 1; Length 319;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 cccaatggaagc 13
 DB 192 CCCAATGGAAGC 181

RESULT 4
 X13773
 ID X13773 standard; DNA; 501 BP.
 AC X13773;
 DT 19-MAR-1999 (first entry)
 DE Enterococcus faecalis genome contig SEQ ID NO:836.
 KW Enterococcus faecalis; contig; detection; Enterococcal infection;
 KW vaccine; attenuation; computer readable medium; ds.
 OS Enterococcus faecalis.
 PN W09850555-A2.
 PD 12-NOV-1998.
 PR 04-MAY-1998; U08985.
 PR 14-NOV-1997; US-066009.
 PR 06-MAY-1997; US-044031.
 PR 16-MAY-1997; US-046635.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Barash SC, Dillon PJ, Kunsch CA;
 PI WPI: 99-045171/04.
 DR New isolated Enterococcus faecalis polynucleotides and polypeptides
 PT used to develop products for the detection of Enterococcus and for
 PT use in vaccines for prevention or attenuation of Enterococcus
 PT infection.
 PS Claim 1: Page 2012; 2084bp; English.
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
 CC X12938 to X13919 represent these nucleotide sequences which are primary
 CC nucleotide sequences, also known as contigs. The computer-based system
 CC can identify fragments of the Enterococcus faecalis genome with
 CC commercial importance. The products can be used to detect the presence
 CC of Enterococcus faecalis in samples. They can also be used for
 CC diagnosing Enterococcal infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to
 CC modulate the growth or pathogenicity of Enterococcus faecalis, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.
 SQ Sequence 501 BP; 162 A; 91 C; 125 G; 117 T;

Query Match 92.3%; Score 12; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagc 12
 DB 408 CCCAATGGAAGC 419

RESULT 5
 O65486
 ID O65486 standard; DNA; 1041 BP.
 AC O65486;
 DT 28-JAN-1995 (first entry)
 DE Gene encoding Elk-L a tyrosine kinase receptor ligand.
 KW Vectors; elk-L protein; elk; ligands; cell growth; differentiation;
 KW ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 signal_peptide 1..72

```

FT      mat_peptide      /*tag- a
FT      73..1038        /*tag- b
PN      W09411384-A.
PD      26-MAY-1994.
PE      15-NOV-1993: U10955.
PR      13-NOV-1992: US-977693.
PA      (IMMUNEX ) IMMUNEX CORP.
PI      Baum PR, Beckmann MP, Lyman S;
DR      WPI: 94-183415/22.
DR      P-PSDB: R55059.
PT      New DNA encoding ligand for elk tyrosine kinase receptor - also
PT      related polypeptides, vectors, antibodies and probes, useful e.g.
PT      in studying cell differentiation or growth
PS      Claim 1: Page 29; 35pp: English.
CC      The sequence is that encoding the elk-L protein able to bind elk, a
CC      tyrosine kinase receptor. The DNA may be incorporated into vectors
CC      which can be used to study the role of elk and its ligands in cell
CC      growth and differentiation.
SQ      Sequence 1041 BP; 239 A; 306 C; 309 G; 187 T;

```

```

Query Match      92.3%; Score 12; DB 1; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ccaatggaagcc 13
DB      413 CCAATGGAAGCC 424

```

```

RESULT 6
T28770 T28770 standard: cDNA to mRNA: 1041 BP.
AC      T28770;
DE      11-DEC-1996 (first entry)
DE      Human cytokine elk-Ligand (elk-L) cDNA.
KW      Human; cytokine; elk-Ligand; elk-L; tyrosine kinase receptor;
KW      neurotrophic; neuroprotective; placenta; radiolabelled probe;
KW      treatment; neural tissue; excitotoxicity; injury; disorder;
KW      neural culture reagent; immunogenic fragment; antibody; ss.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      cds      1..1041
FT      signal_peptide 1..72
FT      /*tag- a
FT      /*tag- b
FT      73..1038
FT      mat_peptide /*tag- c
PN      US5512457-A.
PD      30-APR-1996.
PE      13-NOV-1992: 977693.
PR      13-NOV-1992: US-977693.
PR      15-MAR-1994: US-213403.
PA      (IMMUNEX ) IMMUNEX CORP.
PI      Baum PR, Beckmann MP, Carpenter MK, Lyman S;
DR      WPI: 96-229866/23.
DR      P-PSDB: R91930.
PT      DNA coding for neurotrophic human elk ligand cytokine - useful as
PT      probe to isolate other elk ligand sequences
PS      Claim 7: Columns 27-30; 18pp: English.
CC      The present sequence encodes the human cytokine elk-Ligand (elk-L),
CC      which binds a member of the tyrosine kinase receptor family. Elk-L
CC      exhibits neurotrophic and neuroprotective properties, and has a
CC      calculated mol. wt. 35180 daltons and a pI of 9.006. The elk-L
CC      cDNA, isolated from a human placental cDNA library, can be
CC      radiolabelled and used as a probe for isolating other mammalian
CC      elk-L cDNA. Elk-L can be used to treat neural tissue disorders,
CC      partic. excitotoxicity associated injuries or disorders, and as a
CC      neural culture reagent, while immunogenic fragments of elk-L can be
CC      used to generate specific anti-elk-L antibodies.
SQ      Sequence 1041 BP; 239 A; 306 C; 309 G; 187 T;

```

```

Query Match      92.3%; Score 12; DB 1; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ccaatggaagcc 13
DB      413 CCAATGGAAGCC 424

```

```

RESULT 7
T69766 T69766 standard: cDNA to mRNA: 1041 BP.
AC      T69766;
DE      18-AUG-1997 (first entry)
DE      Human elk ligand cDNA.
KW      Human; elk; ligand; elk-L; cytokine; testing; measuring;
KW      purification; neuroprotection; treatment; diabetic; hereditary;
KW      nutritional; neuropathy; neurodegenerative disease;
KW      tissue culture; ss.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      cds      1..1041
FT      /*tag- a
FT      /*tag- b
FT      /*product- elk_Ligand
FT      signal_peptide 1..72
FT      /*tag- a
FT      /*tag- b
FT      73..1038
FT      mat_peptide /*tag- c
PN      US5627267-A.
PD      06-MAY-1997.
PE      13-NOV-1992: 977693.
PR      13-NOV-1992: US-977693.
PR      15-MAR-1994: US-213403.
PR      01-JUN-1995: US-458077.
PA      (IMMUNEX ) IMMUNEX CORP.
PI      Baum PR, Beckmann MP, Lyman S;
DR      WPI: 97-271366/24.
DR      P-PSDB: W19249.
PT      Human elk ligand protein - for diagnostic or therapeutic use, e.g.
PT      as neuro-protective agent
PS      Claim 7: Columns 27-30; 18pp: English.
CC      The present sequence encodes a human elk ligand (elk-L) protein,
CC      which binds elk, has a calculated molecular weight of 35180 and an
CC      isoelectric point of 9.006. Elk-L is a cytokine that can be used to
CC      test cells for elk expression, measure the biological activity of
CC      elk, purify elk by affinity chromatography and as a neuroprotective
CC      agent to treat diabetic, hereditary and nutritional neuropathies
CC      and neurodegenerative diseases. It may also be added to tissue
CC      cultures to prolong the life of neurons. The elk-L cDNA was
CC      isolated from a human placental cDNA library, and is present as a
CC      cDNA insert in the recombinant vector deposited in strain
CC      ATCC 69085.
SQ      Sequence 1041 BP; 239 A; 306 C; 309 G; 187 T;

```

```

Query Match      92.3%; Score 12; DB 1; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ccaatggaagcc 13
DB      413 CCAATGGAAGCC 424

```

```

RESULT 8
T97976 T97976 standard: cDNA to mRNA: 1041 BP.
AC      T97976;
DE      06-MAR-1998 (first entry)
DE      Human elk-L coding sequence.
KW      Human; elk-L; cytokine; ligand; tyrosine kinase receptor; fusion protein;
KW      extracellular domain; immunoglobulin; neurological disease; ss.

```

OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1041
 FT /tag= a
 FT /product= human elk-L protein
 FT sig_peptide 1..72
 FT /tag= b
 FT mat_peptide 73..1038
 FT /tag= c
 PN US5670625-A.
 PD 23-SEP-1997.
 PF 02-JUN-1995; 460741.
 PR 13-MAR-1994; US-213403.
 PR 13-NOV-1992; US-977693.
 PR 02-JUN-1995; US-460741.
 PA (IMV) IMMUNEX CORP.
 PI Baum PR, Beckmann MP, Lyman S;
 DR WPI: 97-479524/44.
 DR P-PSDB; W36055.
 PT Soluble fusion proteins of human elk-L-lygand and Fc immunoglobulin
 PT fragment - and their dimers and oligomers, useful as
 PS neuro-protectants and neurotrophic agents
 PS Example 3: Columns 27-30; 18pp; English.
 CC This is the nucleotide sequence encoding the human elk-L protein, a new
 CC cytokine that is the ligand for the elk tyrosine kinase receptor. The
 CC extracellular domain of the protein (amino acids 1-213) is used to
 CC generate a fusion protein comprising the Fc polypeptide of the human
 CC immunoglobulin G1 (extending from the hinge region to the C-terminus).
 CC The fusion protein (which has the same activities as the natural elk-L
 CC protein) has neuroprotective and neurotrophic activity so is potentially
 CC useful for treating a wide range of neurological diseases.
 CC Sequence 1041 BP; 239 A; 306 C; 309 G; 187 T;
 SQ

Query Match 92.3%; Score 12; DB 1; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaagcc 13
 |||||
 DB 413 CCAATGGAGGCC 424

RESULT 9
 V15226
 ID V15226 standard; cDNA to mRNA; 1041 BP.
 AC V15226;
 DT 27-MAY-1998 (first entry)
 DE Human elk-L encoding cDNA.
 KW Human; elk-L; elk ligand; cytokine; antibody; cell surface receptor;
 KW tyrosine kinase receptor; neural disease; trophic mechanism; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1041
 FT /tag= a
 FT sig_peptide 1..72
 FT /tag= b
 FT mat_peptide 73..1038
 FT /tag= c
 FT /product= "elk-L"
 PN US5728813-A.
 PD 17-MAR-1998.
 PD 12-NOV-1996; 747240.
 PR 15-MAR-1994; US-213403.
 PR 13-NOV-1992; US-977693.
 PR 02-JUN-1995; US-460741.
 PR 12-NOV-1996; US-747240.
 PA (IMV) IMMUNEX CORP.
 PI Baum PR, Beckmann MP, Lyman S;
 DR WPI: 98-206621/18.
 DR P-PSDB; W44323.
 PT Antibodies to elk ligand polypeptides - bind to tyrosine kinase
 PT receptors, useful for treating neural disease

PS Example 3; Column 27-29; 17pp; English.
 CC The present sequence encodes human elk-L (elk ligand). The present
 CC invention describes antibodies to elk-L polypeptides. The elk-L protein
 CC is known to be neuroprotective exhibiting its effect through a trophic
 CC mechanism. Examples of diseases that may be treated with elk-L include
 CC neuropathies e.g. diabetic, hereditary and nutritional neuropathies,
 CC neurodegenerative diseases and other diseases characterised by loss of
 CC function or degeneration of neurons.
 CC Sequence 1041 BP; 239 A; 306 C; 309 G; 187 T;
 SQ

Query Match 92.3%; Score 12; DB 1; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaagcc 13
 |||||
 DB 413 CCAATGGAGGCC 424

RESULT 10
 V08374
 ID V08374 standard; cDNA to mRNA; 1086 BP.
 AC V08374;
 DT 08-FEB-1999 (first entry)
 DE Cellulose synthase PcsA3 coding sequence.
 KW Cellulose synthase; cotton; ds.
 OS Gossypium hirsutum.
 FH Key Location/Qualifiers
 FT CDS 24..1085
 FT /tag= a
 FT /note= "no stop codon given"
 PN EP-875575-A2.
 PD 04-NOV-1998.
 PF 31-MAR-1998; 302489.
 PR 01-APR-1997; JP-083133.
 PA (NISON) NISSHINO IND INC.
 PI Aotsuka S, Hasegawa O, Hayashi T, Ihara Y;
 DR WPI: 98-559440/48.
 DR P-PSDB; W73310.
 PT New DNA encoding cotton cellulose synthase proteins - useful for
 PT controlling cellulose synthesis in prokaryotic or eukaryotic cells
 PS Example 5; Page 36-37; 51pp; English.
 CC This sequence encodes a cellulose synthase of the invention. The DNA can
 CC be used for controlling cellulose synthesis in a cell by introducing the
 CC DNA into the cell, expressing RNA with a nucleotide sequence homologous
 CC to the DNA or a nucleotide sequence complementary to the DNA. The cells
 CC are prokaryotic cells i.e. acetobacterium and/or eukaryotic cells such as
 CC yeasts i.e. Saccharomyces spp., plant cells (cotton plants) and cultured
 CC cells of mammals etc.
 SQ Sequence 1086 BP; 290 A; 206 C; 296 G; 294 T;
 SQ

Query Match 92.3%; Score 12; DB 1; Length 1086;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaagcc 13
 |||||
 DB 21 CCAATGGAGGCC 32

RESULT 11
 X13928
 ID X13928 standard; DNA; 1516 BP.
 AC X13928;
 DT 31-MAR-1999 (first entry)
 DE H. pylori GHPD 701 gene.
 KW GHPD protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KW peptic ulcer disease; ss.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT CDS 51..1466

```

FT      /tag- a
PN      WO9843478-A1.
PD      08-OCT-1998.
PP      01-APR-1998; 006371.
PR      29-JUL-1997; US-902615.
PR      01-APR-1997; US-833457.
PR      24-JUN-1997; US-881227.
PA      (HUMA-) HUMAN GENOME SCI INC.
PA      (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
PI      Al-Garawi A, Kieantchnou H, Miller C, Oomen RP, Tomb J;
PI      WPI: 98-542293/46.
DR      P-PSDB: W98209.
PT      New isolated Helicobacter polynucleotides - used to develop products
PT      for the diagnosis, prevention and treatment of Helicobacter
PT      infections and gastrointestinal diseases
PS      Claim 1, Page 114-116; 2054pp; English.
CC      This sequence represents a polynucleotide of the invention. It was
CC      isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.
CC      The polypeptides can be used for preventing or treating Helicobacter
CC      infections, and gastroduodenal diseases associated with these
CC      infections, including acute, chronic, and atrophic gastritis, and peptic
CC      ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC      for the production of antibodies. The products can also be used for
CC      detection and diagnosis.
SQ      Sequence 1516 BP; 415 A; 310 C; 332 G; 459 T;

```

```

Query Match      92.3%; Score 12; DB 1; Length 1516;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ccaatggaagcc 13
DB      952 CCAATGGAAGCC 963

```

```

RESULT 12
O20231/c
ID      O20231 standard; DNA: 1575 BP.
AC      O20231;
DT      16-APR-1992 (first entry)
DE      Sequence of the assembled small subunit ADPglucose pyrophosphorylase
DE      gene.
KM      Enzyme; glycogen accumulation; starch biosynthesis; ss.
OS      Potato.
FH      Key      Location/Qualifiers
FT      cds      3..1568
FT      /tag- a
PN      WO9119806-A.
PD      26-DEC-1991.
PP      07-JUN-1991; U04036.
PP      18-JUN-1990; US-539763.
PR      07-JUN-1991; US-709663.
PA      (MONS) MONSANTO CO.
PI      Kishore GM;
PI      WPI: 92-024423/03.
DR      P-PSDB: R20254.
PT      Increasing starch content of plants, esp. potato and tomato - by
PT      introducing into genome a gene for ADP-glucose pyro-phosphorylase
PT      activity, fused to plastid transit gene
PS      Disclosure: Fig 5; 120pp; English.
CC      The inventors claim a method for inc. the starch content of a plant
CC      by inc. its ADP glucose pyrophosphorylase (ADPGPP) activity. The
CC      method uses a recombinant DNA molecule encoding a fusion protein of
CC      amino-terminal plastid transit peptide and ADPGPP. The ADPGPP gene
CC      is the glgC gene of E. coli K12, or its mutant form glg C16 (which
CC      causes accumulation of more glycogen and has higher affinity for ADP
CC      glucose in the absence of an activator). A pref. plastid transit
CC      peptide, designated CTP1, is derived from the ssRUBISCO 1A gene of
CC      Arabidopsis thaliana, and the 3'-untranslated region is pref. that
CC      from the napaline synthase (NOS) gene.
SQ      Sequence 1575 BP; 462 A; 322 C; 365 G; 426 T;

```

```

Query Match      92.3%; Score 12; DB 1; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ccaatggaagcc 13
DB      19 CCAATGGAAGCC 8

```

```

RESULT 13
ID      O80085/c
ID      O80085 standard; CDNA: 1575 BP.
AC      O80085;
DT      03-AUG-1995 (first entry)
DE      ADPGPP small subunit gene.
KM      ADP-glucose-pyrophosphorylase; ADPGPP; glgC; transgenic plant;
KM      potato; Escherichia coli; sweetness; dormancy; crop improvement;
OS      Solanum tuberosum.
FH      Key      Location/Qualifiers
FT      cds      3..1568
FT      /tag- a
PN      WO9428149-A.
PD      08-DEC-1994.
PP      18-MAY-1994; U05275.
PP      28-MAY-1993; US-070155.
PA      (MONS) MONSANTO CO.
PI      Barry GF, Kishore GM, Stark DM, Zalewski JC;
PI      WPI: 95-022820/03.
DR      P-PSDB: R66239.
PT      Transforming potato plants to deregulate ADP glucose
PT      pyro-phosphorylase - improving quality, reducing sugar content
PT      and extending dormancy of stored tubers, also new DNA constructs
PT      for transformation.
PS      Disclosure: Page 48-50; 72pp; English.
CC      A potato tuber ADP-glucose-pyrophosphorylase (ADPGPP) small
CC      subunit gene (given in O80085) was assembled by addition of
CC      sequences from the 1st exon of the genomic clone with a nearly
CC      full-length cDNA of the same gene. The primers given in
CC      O80087-89 were used to: introduce a BglII+NotI site at the ATG
CC      codon to facilitate cloning into E. coli and plant expression
CC      vectors; to introduce a SacI 3' cloning site at the stop codon;
CC      and to remove an internal BglII site. The assembled gene was
CC      expressed in E. coli from the recA promoter in a pRecA-gene10L
CC      cassette. An initiating ATG codon was placed using the primer
CC      given in O80090 to express the mature gene.
SQ      Sequence 1575 BP; 460 A; 324 C; 365 G; 426 T;

```

```

Query Match      92.3%; Score 12; DB 1; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 ccaatggaagcc 13
DB      19 CCAATGGAAGCC 8

```

```

RESULT 14
ID      O85025/c
ID      O85025 standard; CDNA: 1575 BP.
AC      O85025;
DT      23-SEP-1995 (first entry)
DE      ADP-glucose-pyrophosphorylase small subunit gene.
KM      Potato tuber; ADP-glucose-pyrophosphorylase small subunit;
KM      site-directed mutagenesis; Escherichia coli; transgenic plant;
KM      enzyme engineering; cloning; crop improvement; oilseed;
KM      feedstuff; food; ds.
OS      Solanum tuberosum.
FH      Key      Location/Qualifiers
FT      cds      3..1568
FT      /tag- a

```

```

FT      /product= ADP-glucose-pyrophosphorylase
FT      /note= "EC-2.7.7.27"
FT      primer_bind
FT      1. .14
FT      b
FT      /note= "site for BglII-NcoI site introduction at
FT      ATG codon (primer Q85016)"
FT      complement (1555. .1575)
FT      primer_bind
FT      /tag= c
FT      /note= "site for SacI site insertion at stop codon
FT      (primer Q85017)"
FT      complement (406. .435)
FT      primer_bind
FT      /tag= d
FT      /note= "site for internal BglII site removal
FT      (primer Q85018)"
FT      complement (234. .435)
FT      primer_bind
FT      /tag= e
FT      /note= "site for Met initiation codon insertion
FT      (primer Q85019)"
FT      EP-634491-A.
PN      18-JAN-1995.
PD      11-JUL-1994; 870118.
PR      12-JUL-1993; US-090523.
PA      (MONS ) MONSANTO CO.
PI      Barry GF, Kishore GM, Stark DM;
DR      WPI; 95-045180/07.
DR      P-PSDB; R68972.
PT      Prodn of seeds e.g. peanut or soybean having decreased oil
PT      content - by providing increased levels of ADP glucose
PT      pyro:phosphorylase
PS      Disclosure; Page 21-24; 35pp; English.
CC      The sequence represents a potato tuber ADP-glucose-pyrophosphorylase
CC      small subunit gene. The gene is modified by site-directed
CC      mutagenesis to include a BglII-NcoI site at the ATG codon, a SacI
CC      site at the stop codon, and an internal BglII site is removed. A
CC      methionine initiation codon is introduced for expression of the
CC      mature protein in Escherichia coli. The gene is cloned together
CC      with the large subunit gene (Q94439) (on compatible plasmids), and
CC      subjected to mutagenesis to give improved variants of the enzyme.
CC      The resulting variant genes may be expressed in an oilseed crop.
CC      plant to reduce the oil content in seeds, for use in feedstuff and
CC      food products.
SQ      Sequence 1575 BP; 461 A; 323 C; 365 G; 426 T;

Query Match          92.3%; Score 12; DB 1; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 ccaatggaagcc 13
DB      19 CCAATGGAAGCC 8

RESULT 15
T29734
ID      T29734 standard; DNA: 1836 BP.
AC      T29734;
DT      27-NOV-1996 (first entry)
DE      Brevibacterium flavum MJ-233 dihydroxy acid dehydratase coding region.
KW      Dihydroxy-acid dehydratase; EC 4.2.1.9; amino acid biosynthesis;
KW      Isoleucine; Ile; valine; Val; Corynebacterium bacterium; ds.
OS      Brevibacterium flavum MJ-233.
FH      Key Location/Qualifiers
FT      met_peptide
FT      1. .1836
FT      /tag= a
FT      /EC_number= 4.2.1.9
FT      J08089249-A.
PD      09-APR-1996.
PR      29-SEP-1994; 234612.
PR      29-SEP-1994; JP-234612.
PA      (MITU ) MITSUBISHI CHEM CORP.
DR      WPI; 96-233342/24.
DR      P-PSDB; R91947.

```

```

PT      DNA fragment contg. a gene coding for dihydroxy-acid dehydratase -
PT      derived from Brevibacterium flavum MJ 233, useful for prodn. of
PT      L-Isoleucine and L-valine
PS      Claim 5; Page 9-11; 12pp; Japanese.
CC      A 6 kb genomic DNA KpnI fragment containing an open reading frame
CC      (ORF) coding for dihydroxy-acid dehydratase was isolated from
CC      Brevibacterium flavum MJ-233. The sequence of the 1836 bp ORF was
CC      determined (i.e. the present sequence) and was found to encode a
CC      protein of 612 amino acids. The enzyme is involved in biosynthesis
CC      of the amino acids Isoleucine and valine.
SQ      Sequence 1836 BP; 426 A; 606 C; 475 G; 329 T;

Query Match          92.3%; Score 12; DB 1; Length 1836;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ccaatggaagc 12
DB      438 CCAATGGAAGC 449

Search completed: September 12, 2000, 23:07:26
Job time: 4108 sec

```

Wed Sep 13 11:14:11 2000

us-09-542-718-5.rng

Page 7

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 23:04:39 ; Search time 112.94 Seconds

(without alignments)
15.833 Million cell updates/sec

Title: US-09-542-718-5

Perfect score: 13

Sequence: 1 cccaatggaagcc 13

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 243080 seqs, 68777915 residues

Total number of hits satisfying chosen parameters: 486160

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/5C_COMB.seq:*
4: /cgn2_6/prodata/1/lna/5D_COMB.seq:*
5: /cgn2_6/prodata/1/lna/5E_COMB.seq:*
6: /cgn2_6/prodata/1/lna/PCTUS_COMB.seq:*
7: /cgn2_6/prodata/1/lna/backfileseq1.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	12	92.3	319	2	US-08-592-126-126
2	12	92.3	1041	1	US-08-213-403-1
3	12	92.3	1041	1	US-08-458-077-1
4	12	92.3	1041	1	US-08-460-741-1
5	12	92.3	1041	2	US-08-747-240-1
6	12	92.3	1575	1	US-08-090-523-7
7	12	92.3	1575	1	US-08-398-627-7
8	12	92.3	1575	1	US-08-406-858-7
9	12	92.3	1575	6	PCT-US91-04036-7
10	12	92.3	1575	6	PCT-US94-05275-7
11	12	92.3	2015	1	US-07-912-015-1
12	12	92.3	2625	3	US-08-468-036-2
13	12	92.3	2625	3	US-08-376-843-2
14	12	92.3	21126	1	US-08-008-216-19
15	12	92.3	21126	1	US-08-459-569-19
16	12	92.3	21126	2	US-08-458-831-19
17	12	92.3	20	1	US-08-049-061-6
18	12	92.3	20	1	US-08-049-061-10
19	12	92.3	20	1	US-08-049-061-11
20	12	92.3	20	1	US-08-049-061-27
21	12	92.3	20	1	US-08-049-061-28
22	12	92.3	22	1	US-08-049-061-9
23	12	92.3	24	1	US-08-049-061-12
24	12	92.3	26	1	US-08-049-061-13
25	12	92.3	28	1	US-08-049-061-14
26	12	92.3	40	1	US-08-049-061-22

27	11.4	87.7	40	1	US-08-049-061-23	Sequence 23, Appl
28	11.4	87.7	270	2	US-08-594-031-109	Sequence 109, App
29	11.4	87.7	309	2	US-08-594-031-106	Sequence 106, App
30	11.4	87.7	357	3	US-08-294-143-1	Sequence 1, Appl
31	11.4	87.7	473	2	US-08-594-031-115	Sequence 115, App
32	11.4	87.7	521	2	US-08-463-115-10	Sequence 10, Appl
33	11.4	87.7	521	2	US-08-465-388-10	Sequence 10, Appl
34	11.4	87.7	524	2	US-08-594-031-105	Sequence 105, App
35	11.4	87.7	564	1	US-08-136-993-11	Sequence 11, App
36	11.4	87.7	620	2	US-08-594-031-113	Sequence 113, App
37	11.4	87.7	645	1	US-08-312-870-4	Sequence 4, Appl
38	11.4	87.7	829	3	US-08-294-143-3	Sequence 3, Appl
39	11.4	87.7	878	2	US-08-469-667-8	Sequence 8, Appl
40	11.4	87.7	878	6	PCT-US95-07289-8	Sequence 8, Appl
41	11.4	87.7	993	1	US-08-136-993-3	Sequence 3, Appl
42	11.4	87.7	1055	3	US-08-468-012A-1	Sequence 1, Appl
43	11.4	87.7	1140	3	US-08-704-548-1	Sequence 1, Appl
44	11.4	87.7	1147	2	US-08-417-103-15	Sequence 15, Appl
45	11.4	87.7	1218	1	US-08-136-993-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-592-126-126/C
; Sequence 126, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,126
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 126:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: G30.seq
; US-08-592-126-126

Query Match 92.3%; Score 12; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaacc 13
|||||
Db 192 CCAATGGAAGCC 181

RESULT 2

US-08-213-403-1
: Sequence 1, Application US/08213403
: Patent No. 5512457
: GENERAL INFORMATION:
: APPLICANT: Lyman, Stewart D.
: APPLICANT: Beckmann, M. Patricia
: APPLICANT: Baum, Peter R
: APPLICANT: Carpenter, Melissa
: TITLE OF INVENTION: No. 5512457e1 Cytokine Designated elk ligand
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: Apple Macintosh
: SOFTWARE: Microsoft Word for Apple, Version 5.1a
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/213,403
: FILING DATE: 15-MAR-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/977,693
: FILING DATE: 13-NOV-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Seese, Kathryn A.
: REGISTRATION NUMBER: 32,172
: REFERENCE/DOCKET NUMBER: 2807-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-0644
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1041 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1041
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 73..1038
US-08-213-403-1

Query Match 92.3%; Score 12; DB 1; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaacc 13
|||||
Db 413 CCAATGGAAGCC 424

RESULT 3

US-08-458-077-1
: Sequence 1, Application US/08458077

: Patent No. 5627267
: GENERAL INFORMATION:
: APPLICANT: Lyman, Stewart D.
: APPLICANT: Beckmann, M. Patricia
: APPLICANT: Baum, Peter R
: APPLICANT: Carpenter, Melissa
: TITLE OF INVENTION: No. 5627267e1 Cytokine Designated elk ligand
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: Apple Macintosh
: SOFTWARE: Microsoft Word for Apple, Version 5.1a
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/458,077
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/213,403
: FILING DATE: 15-MAR-1994
: APPLICATION NUMBER: US 07/977,693
: FILING DATE: 13-NOV-1992
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Seese, Kathryn A.
: REGISTRATION NUMBER: 32,172
: REFERENCE/DOCKET NUMBER: 2807-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-0644
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1041 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1041
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 73..1038
US-08-458-077-1

Query Match 92.3%; Score 12; DB 1; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaacc 13
|||||
Db 413 CCAATGGAAGCC 424

RESULT 4

US-08-460-741-1
: Sequence 1, Application US/08460741
: Patent No. 5670625
: GENERAL INFORMATION:
: APPLICANT: Lyman, Stewart D.
: APPLICANT: Beckmann, M. Patricia
: APPLICANT: Baum, Peter R
: APPLICANT: Carpenter, Melissa

```

; TITLE OF INVENTION: No. 5670625el Cytokine Designated elk Ligand
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,741
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; APPLICATION NUMBER: US 08/213,403
; FILING DATE: 15-MAR-1994
; APPLICATION NUMBER: US 07/977,693
; FILING DATE: 13-NOV-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2807-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1041
; NAME/KEY: mat_peptide
; LOCATION: 73..1038
; US-08-460-741-1

Query Match          92.3%: Score 12; DB 1; Length 1041;
Best Local Similarity 100.0%: Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      2 ccaatggaacc 13
Db      413 CCAATGGAACC 424

RESULT 5
; Sequence 1, Application US/08747240
; Patent No. 5728813
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Beckmann, M. Patricia
; APPLICANT: Baum, Peter R.
; APPLICANT: Carpenter, Melissa
; TITLE OF INVENTION: No. 5728813el Cytokine Designated elk Ligand
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
```

```

; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,240
; FILING DATE: 12-NOV-1996
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 08/213,403
; FILING DATE: 15-MAR-1994
; APPLICATION NUMBER: US 07/977,693
; FILING DATE: 13-NOV-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seese, Kathryn A.
; REGISTRATION NUMBER: 32,172
; REFERENCE/DOCKET NUMBER: 2807-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1041 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1041
; NAME/KEY: mat_peptide
; LOCATION: 73..1038
; US-08-747-240-1

Query Match          92.3%: Score 12; DB 2; Length 1041;
Best Local Similarity 100.0%: Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      2 ccaatggaacc 13
Db      413 CCAATGGAACC 424

RESULT 6
; Sequence 7, Application US/08090523
; Patent No. 5498830
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Stark, David M.
; TITLE OF INVENTION: Enhanced Starch Biosynthesis
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Co. BB4F
; STREET: 700 Chesterfield Parkway No. 5498830th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/090,523
FILING DATE: 19930712
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/709663
FILING DATE: 07-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539763
FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10559)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-6047
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1565
US-08-090-523-7

Query Match 92.3%; Score 12; DB 1; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaagcc 13
Db 19 CCAATGGAAGCC 8
RESULT 7
US-08-398-627-7/C
Sequence 7, Application US/08398627
Patent No. 5608149
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Stark, David M.
TITLE OF INVENTION: Enhanced Starch Biosynthesis
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Co. BBAF
STREET: 700 Chesterfield Parkway No. 5608149th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,627
FILING DATE: 03-MAR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/090,523
FILING DATE: 12-JUL-1993
APPLICATION NUMBER: US 07/709663
FILING DATE: 07-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/539763

FILING DATE: 18-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10559)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-7286
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1565
US-08-398-627-7

Query Match 92.3%; Score 12; DB 1; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaagcc 13
Db 19 CCAATGGAAGCC 8
RESULT 8
US-08-406-858-7/C
Sequence 7, Application US/08406858
Patent No. 5648249
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Stark, David M.
APPLICANT: Zaleski, James C.
TITLE OF INVENTION: Method of improving the quality of
STARCHES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grace L. Bonner, Monsanto Company, BBAF
STREET: 700 Chesterfield Parkway No. 5648249th
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/406,858
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05275
FILING DATE: 18-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/070,155
FILING DATE: 28-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bonner, Grace L.
REGISTRATION NUMBER: 32,963
REFERENCE/DOCKET NUMBER: 38-21(10654)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6047
TELEFAX: (314)537-7286
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:

LENGTH: 1575 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1565
US-08-406-858-7

Query Match 92.3%; Score 12; DB 1; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ccaatggaagcc 13
|||||

DB 19 CCAATGGAAGCC 8

RESULT 9
PCT-US91-04036-7/C
Sequence 7, Application PC/TUS9104036
GENERAL INFORMATION:
APPLICANT: Kishore, Ganesh M.
TITLE OF INVENTION: Increased Starch Content In Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Monsanto Co.
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/04036
FILING DATE: 19910607
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: McBride, Thomas P.
REGISTRATION/DOCKET NUMBER: 32706
REFERENCE/DOCKET NUMBER: 38-21(10530)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 537-7357
TELEFAX: (314) 537-6047
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1565
PCT-US91-04036-7

QY 2 ccaatggaagcc 13
|||||

DB 19 CCAATGGAAGCC 8

Query Match 92.3%; Score 12; DB 6; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ccaatggaagcc 13
|||||

DB 19 CCAATGGAAGCC 8

RESULT 10

PCT-US94-05275-7/C
Sequence 7, Application PC/TUS9405275
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Method of Improving the Quality of Stored
TITLE OF INVENTION: Potatoes
NUMBER OF SEQUENCES: 26
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05275
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/070155
FILING DATE: 28-MAY-1993
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1565
PCT-US94-05275-7

Query Match 92.3%; Score 12; DB 6; Length 1575;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ccaatggaagcc 13
|||||

DB 19 CCAATGGAAGCC 8

RESULT 11
US-07-912-015-1/C
Sequence 1, Application US/07912015
Patent No. 5283191
GENERAL INFORMATION:
APPLICANT: Morgan, Robin Wilson
APPLICANT: Claessens, Johannes Antonius Joseph
APPLICANT: Sondermeijer, Paulus Jacobus Antonius
TITLE OF INVENTION: Marek's Disease Virus Vaccine
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Organon Teknika Corporation
ADDRESS: Biotechnology Research Institute
STREET: 1330-A Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/912,015
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/615,211
FILING DATE: 16-NOV-1990
ATTORNEY/AGENT INFORMATION:

```
NAME: Donna Bobrowicz
REGISTRATION NUMBER: 32,196
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301)258-5200
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2015 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Marek's disease herpesvirus
STRAIN: GA
FEATURE:
NAME/KEY: CDS
LOCATION: 14..2005
OTHER INFORMATION: /label= pMD18
US-07-912-015-1

Query Match          92.3%; Score 12; DB 1; Length 2015;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cccaatggaac 12
    |||
Db 1224 CCCAATGGAAC 1213

RESULT 12
US-08-468-036-2
; Sequence 2, Application US/08468036
; Patent No. 5728806
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins that
; TITLE OF INVENTION: Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,036
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5728806and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2625 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 796..2580
; US-08-376-843-2
```

```
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 796..2580
US-08-468-036-2

Query Match          92.3%; Score 12; DB 2; Length 2625;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cccaatggaac 12
    |||
Db 495 CCCAATGGAAC 506

RESULT 13
US-08-376-843-2
; Sequence 2, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; TITLE OF INVENTION: that Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/376,843
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5846764and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2625 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 796..2580
; US-08-376-843-2

Query Match          92.3%; Score 12; DB 3; Length 2625;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cccaatggaac 12
    |||
Db 495 CCCAATGGAAC 506
```

```

RESULT 14
US-08-008-216-19
; Sequence 19, Application US/08008216
; Patent No. 5366887
; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Tepfer, David A.
; TITLE OF INVENTION: RI T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY, CARY, AMES & FRYE
; STREET: 401 B Street, Suite 1700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-4297
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,216
; FILING DATE: 25-JAN-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/725,368
; FILING DATE: 22-APR-1985
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhorst, Marile W.
; REGISTRATION NUMBER: 36,740
; REFERENCE/DOCKET NUMBER: P1020US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 699-2700
; TELEFAX: (619) 236-1048
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Agrobacterium rhizogenes
; STRAIN: STRAIN A4
; IMMEDIATE SOURCE:
; LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
; CLONE: CLONE 7
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (937..2262)
; OTHER INFORMATION: /label= ORF1SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (2649..3458)
; OTHER INFORMATION: /label= ORF2SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3726..4799
; OTHER INFORMATION: /label= ORF3SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (4041..4400)
; OTHER INFORMATION: /label= ORF4SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (4607..4918)
; OTHER INFORMATION: /label= ORF5SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: 5143..6216
; OTHER INFORMATION: /label= ORF6SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (5071..5643)
; OTHER INFORMATION: /label= ORF7SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6609..8888
; OTHER INFORMATION: /label= ORF8SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (6576..6830)
; OTHER INFORMATION: /label= ORF9SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9748..10044
; OTHER INFORMATION: /label= ORF10SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (10509..11282)
; OTHER INFORMATION: /label= ORF11SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12466..13002
; OTHER INFORMATION: /label= ORF12SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 13723..14319
; OTHER INFORMATION: /label= ORF13SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15659..16210
; OTHER INFORMATION: /label= ORF14SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (16517..17545)
; OTHER INFORMATION: /label= ORF15SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (17737..18189)
; OTHER INFORMATION: /label= ORF16SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (18177..18743)
; OTHER INFORMATION: /label= ORF17SUBSEQUENCE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: complement (19031..19390)
; OTHER INFORMATION: /label= ORF18SUBSEQUENCE
; US-08-008-216-19

Query Match          92.3%; Score 12; DB 1; Length 21126;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ccaatggaagcc 13
      |||
Db 19549 CCAATGGAAGCC 19560

RESULT 15
US-08-459-569-19
; Sequence 19, Application US/08459569
; Patent No. 5543501
; GENERAL INFORMATION:
; APPLICANT: Slightom, Jerry L.
; APPLICANT: Tepfer, David A.
; TITLE OF INVENTION: RI T-DNA Promoters
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GRAY, CARY, AMES & FRYE

```

```

STREET: 401 B Street, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-4297
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,569
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,216
FILING DATE: 25-JAN-1993
APPLICATION NUMBER: US 06/725,368
FILING DATE: 22-APR-1985
ATTORNEY/AGENT INFORMATION:
NAME: Barthorst, Marnie W.
REGISTRATION NUMBER: 36,740
REFERENCE/DOCKET NUMBER: P1020U51
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-2700
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 21126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Agrobacterium rhizogenes
STRAIN: STRAIN A4
IMMEDIATE SOURCE:
LIBRARY: CONVOLVULUS ARVENSIS PLANT CELLS
CLONE: CLONE 7
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (937..2262)
OTHER INFORMATION: /label= ORF1SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (2649..3458)
OTHER INFORMATION: /label= ORF2SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3726..4799
OTHER INFORMATION: /label= ORF3SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (4041..4400)
OTHER INFORMATION: /label= ORF4SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (4607..4918)
OTHER INFORMATION: /label= ORF5SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5143..6216
OTHER INFORMATION: /label= ORF6SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (5071..5643)
OTHER INFORMATION: /label= ORF7SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6609..8888
OTHER INFORMATION: /label= ORF8SUBSEQUENCE
FEATURE:

```

```

NAME/KEY: misc_feature
LOCATION: complement (6576..6830)
OTHER INFORMATION: /label= ORF9SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9748..10044
OTHER INFORMATION: /label= ORF10SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (10509..11282)
OTHER INFORMATION: /label= ORF11SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12466..13002
OTHER INFORMATION: /label= ORF12SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13723..14319
OTHER INFORMATION: /label= ORF13SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15659..16210
OTHER INFORMATION: /label= ORF14SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (17737..18189)
OTHER INFORMATION: /label= ORF15SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (18177..18743)
OTHER INFORMATION: /label= ORF17SUBSEQUENCE
FEATURE:
NAME/KEY: misc_feature
LOCATION: complement (19031..19390)
OTHER INFORMATION: /label= ORF18SUBSEQUENCE
US-08-459-569-19

```

```

Query Match          92.3%; Score 12; DB 1; Length 21126;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2 ccaatggaagcc 13
        |||
Db 19549 CCAATGGAAGCC 19560

```

Search completed: September 12, 2000, 23:04:42
Job time: 3949 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2000, 22:46:55 : Search time 1893.64 Seconds
(without alignments)
30.274 Million cell updates/sec

Title: US-09-542-718-5

Perfect score: 13

Sequence: 1 cccaatggaagcc 13

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5247842 seqs, 2204914090 residues

Total number of hits satisfying chosen parameters: 10495684

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: em_est35:*
83: em_est36:*
84: em_est37:*
85: gb_est48:*
86: gb_est49:*
87: gb_est50:*
88: gb_est51:*
89: gb_est52:*
90: gb_est53:*
91: gb_est54:*
92: gb_est55:*
93: gb_est56:*
94: gb_est57:*
95: gb_est58:*
96: gb_est59:*
97: em_est38:*
98: em_est39:*
99: em_est40:*
100: em_est41:*
101: gb_est60:*
102: gb_est61:*
103: gb_est62:*
104: gb_est63:*
105: gb_est64:*
106: em_est42:*
107: em_est43:*
108: em_est44:*
109: em_est45:*
110: em_est46:*
111: em_est47:*
112: em_est48:*
113: gb_est65:*
114: gb_est66:*
115: em_est49:*
116: gb_est67:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	100.0	164	71	AM396623 sg80d10.y
2	13	100.0	175	38	AI379962 tc81b06.x
3	13	100.0	189	60	AV281774 AV281774
4	13	100.0	199	71	AM324305 uc027a01.x
5	13	100.0	215	62	AV386986 AV386986
6	13	100.0	236	59	AV278622 AV278622
7	13	100.0	254	23	AA282475 z590c01.s
8	13	100.0	293	80	AM698136 NKNV.073
9	13	100.0	294	86	N28122 M081422R.MO
10	13	100.0	326	38	AI352128 q110f07.x
11	13	100.0	332	72	AM411422 fh12c03.y
12	13	100.0	338	31	AA786244 j7h11a1.f
13	13	100.0	348	45	AI845473 UI-M-X01
14	13	100.0	348	44	AI043778 UI-R-C0-1
15	13	100.0	350	89	T00893 WEST01614.E
16	13	100.0	358	74	AM597076 sj69h12.y
17	13	100.0	360	81	D69118 CELK063A6F
18	13	100.0	360	81	D69385 CELK063E9F
19	13	100.0	360	81	D69742 CELK073EYF
20	13	100.0	361	31	AA785620 q9f05a1.f
21	13	100.0	365	62	AM025820 wu08c09.x
22	13	100.0	366	63	AM028366 w788b03.x
23	13	100.0	368	120	B63027 C1T978SK-13
24	13	100.0	371	72	AM442811 EST307741
25	13	100.0	374	33	AA966772 sg01a1.f
26	13	100.0	376	34	AI022536 ow58b03.s
27	13	100.0	379	28	AA570786 nm37e03.s
28	13	100.0	379	119	AA2037328 RPT-23-3
29	13	100.0	380	88	R19461 y925h04.r1
30	13	100.0	384	79	AM676321 833002H10
31	13	100.0	384	88	R17042 yf45f05.r2
32	13	100.0	387	31	AA788418 r7h07a1.f
33	13	100.0	390	29	AA639559 ng91f08.s
34	13	100.0	395	37	AI273585 q160a11.x
35	13	100.0	406	35	AI093637 ou83d11.s
36	13	100.0	407	101	AQ362888 nbx00051M
37	13	100.0	413	95	AQ208746 HS-3229.A
38	13	100.0	414	34	AI050032 an31a04.x
39	13	100.0	417	64	AM136253 UI-H-B11
40	13	100.0	420	64	AM137518 UI-H-B11
41	13	100.0	423	93	AQ019463 C1T-HSP-2
42	13	100.0	428	38	AI367860 q945f05.x
43	13	100.0	430	35	AI123505 qe01g08.x
44	13	100.0	433	116	AQ863876 nbep0022K
45	13	100.0	443	35	AI092651 q88b003.x

ALIGNMENTS

RESULT 1
LOCUS AM396623 164 bp mRNA
DEFINITION sg80d10.y1 Gm-c1026 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1026-44 5' similar to TR:080828 080828 HYPOTHETICAL 88.8 KD

PROTEIN. ; mRNA sequence.

ACCESSION AM396623
VERSION AM396623.1 GI:6915162
KEYWORDS EST.

SOURCE soybean.

ORGANISM Glycine max

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

1 (bases 1 to 164)

Shoemaker,R., Keim,P., Vockin,L., Erpelidg,J., Corryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

REFERENCE

AUTHORS

Public Soybean EST Project

Unpublished (1999)

On Jan 6, 2000 this sequence version replaced gi:6675643.

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information

call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or

info@genomesystems.com web site: www.genomesystems.com

possible reversed clone: similarity on wrong strand.

location/Qualifiers

source

1..164

/organism="Glycine max"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1026-44"

/clone_lib="Gm-c1026"

/tissue_type="Senescing leaves, mature plants, greenhouse grown."

/lab_host="DH10B"

/note="Vector: pRT73-Pac (Pharmacia); Site_1: EcoRI; Site_2: HindIII; This cDNA library was constructed from

mRNA isolated from senescing leave tissue of mature

greenhouse grown plants. Complementary DNA was synthesized

from mRNA using a 3' anchored poly(dT) primer. EcoRI

adapters were ligated to the blunt-ended cDNA fragments

followed by digestion with EcoRI and HindIII. The cDNA

fragments were directionally cloned into the EcoRI-HindIII

restriction site of the pRT73-Pac vector. The ligated cDNA

fragments were transformed into DH10B host cells (Gibco

BR). This library was constructed R. Shoemaker and J.

Erpelidg."

BASE COUNT 47 a 31 c 37 g 49 t

ORIGIN

Query Match 100.0%; Score 13; DB 71; Length 164;

Best Local Similarity 100.0%; Pred. No. 7.8e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cccaatgaagcc 13

|||||

Db 66 CCCAATGAGGCC 54

RESULT 2

LOCUS AI379962

DEFINITION tc81b06.x1 NCI-GCAP-CL11 Homo sapiens cDNA clone IMAGE:2072531 3'

similar to SW:02RP_HUMAN P05161 UBIOUITIN CROSS-REACTIVE PROTEIN ;

ACCESSION AI379962

```

VERSION      AI379962.1   GI:4189815
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 175)
AUTHORS      NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT       On May 18, 1998 this sequence version replaced gi:3137165.
             Contact: Robert Strausberg, Ph.D.
             Tel.: (301) 496-1550
             Email: Robert.Strausberg@nih.gov
             Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
             M.D., Louis M. Staudt, M.D., Ph.D.
             CDNA Library Preparation: M. Bento Soares, Ph.D.
             CDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LNL at:
             www.bio.lnl.gov/bdrrp/image/image.html

FEATURES
  source
    1..175
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="IMAGE:2072531"
       /clone_id="NCI_CGAP_CLU1"
       /tissue_type="B-cell, chronic lymphocytic leukemia"
       /lab_host="DH10B"
       /note="Vector: pT7TS-Pec (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTCGGAGCGCGCATTCGTCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7TS vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldó."

BASE COUNT
  47 a      48 c      34 g      46 t

ORIGIN
Query Match          100.0%; Score 13; DB 38; Length 175;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 cccaatggaagcc 13
         |||
Db       45 CCCAATGGAAGC 57

RESULT      3
AVZ81774    LOCUS           189 bp      mRNA      EST      05-NOV-1999
DEFINITION AVZ81774 RIKEN full-length enriched, adult male testis (DH10B) Mus
            musculus cDNA clone 4933426F15 3', mRNA sequence.
ACCESSION  AVZ81774
VERSION     AVZ81774.1  GI:6269811
KEYWORDS    EST.
SOURCE      house mouse.
            Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 189)
AUTHORS      Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
            Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,

```

TITLE
 JOURNAL
 COMMENT

Iishi, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I.,
 Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M.,
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
 Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K.,
 Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugihara, Y.,
 Suzuki, H., Suzuki, H., Takehashi, F., Tateo, N., Tomimaga, N.,
 Tsunoda, T., Watanabe, S., Yamamura, T., Yasunishi, A.,
 Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Kono, H., et al.)
 Unpublished (1999)
 On Apr 7, 1998 this sequence version replaced gi:3034900.
 Contact: Yoshihide Hayashizaki
 Genome Exploration Research Group, Life Science Tsukuba Center,
 The Institute of Physical and Chemical Research (RIKEN), Genomic
 Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: +81-298-36-9013
 Fax: +81-298-36-9098
 Email: genome-res@rtc.riken.go.jp,
 URL: http://genome.rtc.riken.go.jp/
 Sasaki, N., Izawa, M., Wataniki, M., Ozawa, K., Tanaka, T., Yoneda, Y.,
 Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and
 Hayashizaki, Y.
 Transcriptional sequencing: A method for DNA sequencing using RNA
 polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
 Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.
 Location/Organisms
 1. 189
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4933426F15"
 (clone_id="RIKEN full-length enriched, adult male testis
 (DH10B)")
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site_1: Sali; Site_2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGGAGAGAGATCTCGAGATCTTTTAAATTTAAATCCGCCGCCGCC 3']. cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adapter of sequence [5'
 GAGGAGAGATCTCGAGATCTTTTAAATTTAAATCCGCCGCCGCC 3']. cDNA
 was cloned into the XhoI and BamHI sites. Vector: a
 modified pBluescript KS(+) after bulk excision from Lambda
 FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI."

BASE COUNT
 ORIGIN

65 a 42 c 25 g 57 t

Query Match 100.0%; Score 13; DB 60; Length 189;
 Best Local Similarity 100.0%; Pred. No. 8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 cccaatggaagcc 13

Db 3 CCCAATGGAAGCC 15

RESULT 4
LOCUS AM324305
DEFINITION u027a01.x1 NCI_CGAP_Mam6 Mus musculus cdna clone IMAGE:2616072 3', mRNA sequence.
ACCESSION AM324305
VERSION AM324305.1
KEYWORDS GI:6758330
SOURCE EST.
ORGANISM house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 199)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT Unpublished (1997)
Other ESTs: u027a01.y1
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Library distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/dbtp/image/image.html

BASE COUNT
MGI:1024172
Seq primer: -40UP from Gibco
High quality sequence stop: 191.

FEATURES
Location/Qualifiers
1..199
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:2616072"
/clone_1ib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue.type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: Not; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT
55 a 40 c 54 g 50 t

QUERY MATCH
Best Local Similarity 100.0%; Score 13; DB 71; Length 199;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cccaatggaagcc 13
Db 26 CCCAATGGAAGCC 38

RESULT 5
LOCUS AV386986
DEFINITION AV386986 Chlamydomonas reinhardtii C9 Chlamydomonas reinhardtii
CDNA clone CM010g11_r, mRNA sequence.
ACCESSION AV386986
VERSION AV386986.1
KEYWORDS GI:6541202
SOURCE Chlamydomonas reinhardtii.

ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE
AUTHORS 1 (bases 1 to 215)
TITLE Asamizu, E., Nakamura, Y., Sato, S., Fukuzawa, H. and Tabata, S.
A large scale structural analysis of cDNAs in a unicellular green Alga, Chlamydomonas reinhardtii. I. Generation of 3433 Non-redundant Expressed Sequence Tags
DNA Res. 6, 369-373 (1999)
On May 18, 1998 this sequence version replaced gi:3138387.
Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..215
/organism="Chlamydomonas reinhardtii"
/strain="C9"
/db_xref="taxon:3055"
/clone="CM010g11_r"
/clone_1ib="Chlamydomonas reinhardtii C9"
/dev_stage="photoautotrophic growth"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT
53 a 70 c 60 g 32 t

QUERY MATCH
Best Local Similarity 100.0%; Score 13; DB 62; Length 215;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cccaatggaagcc 13
Db 143 CCCAATGGAAGCC 155

RESULT 6
LOCUS AV278622
DEFINITION AV278622 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus cdna clone 4933402018 3', mRNA sequence.
ACCESSION AV278622
VERSION AV278622.1
KEYWORDS GI:6266659
SOURCE EST.
ORGANISM house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 236)
Kono, H., Alizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, Y., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, R., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomiwa, N., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al.)
Unpublished (1999)
On Jul 8, 1999 this sequence version replaced gi:5420947.
Contact: Yoshihide Hayashizaki
Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9098
Email: genome-res@rcc.riken.go.jp,

URL: <http://genome.rtc.riken.go.jp/>
 Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 95 (7), 3455-3460 (1998)
 Itoh, M., Kikunishi, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation system. *Genome Res.* 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. *Methods Enzymol.* 303, 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp/>) for further details.

FEATURES

source

Location/Qualifiers
 1..236
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4934402018"
 /clone_1lb="RIKEN full-length enriched, adult male testis (DH10B)"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
 GAGAGAGAGATCGATCGAAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGATCGATCGATTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pUC19 after bulk excision from Lambda FIC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
 BASE COUNT 73 a 56 c 44 g 63 t
 ORIGIN

Query Match 100.0%; Score 13; DB 59; Length 236;
 Best Local Similarity 100.0%; Pred. No. 8.3e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
 |||||
 Db 49 CCCAATGGAAGCC 61

RESULT 7
 AA282475 254 bp mRNA EST 13-AUG-1997
 LOCUS AA282475/c 289001.81 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:704736 3',
 DEFINITION mRNA sequence.
 ACCSSION AA282475
 VERSION AA282475.1 GI:1925437
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 254)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT On Jan 14, 1998 this sequence version replaced gi:1798137.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNLN; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1105 Std Error: 0.00
 Seq primer: -41m13 fwd. ET from Amerisham
 High quality sequence stop: 244.

FEATURES

source

Location/Qualifiers
 1..254
 /organism="Homo sapiens"
 /db_xref="GDB:5654460"
 /db_xref="taxon:9606"
 /clone="IMAGE:704736"
 /clone_1lb="NCI_CGAP_GCB1"
 /tissue_type="germlinal center B cell"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, 19D-), provided by Dr. Gerald Merl (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTCACATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 31 a 70 c 83 g 70 t
 ORIGIN

Query Match 100.0%; Score 13; DB 23; Length 254;
 Best Local Similarity 100.0%; Pred. No. 8.4e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
 |||||
 Db 175 CCCAATGGAAGCC 163

RESULT 8
 AM698136 293 bp mRNA EST 17-APR-2000
 LOCUS NXNV_073_C07_F Nsf Xylem Normal wood Vertical Pinus taeda CDNA
 DEFINITION clone NXNV_073_C07 5', mRNA sequence.
 ACCSSION AM698136
 VERSION AM698136.1 GI:7580722
 KEYWORDS EST.
 SOURCE loblolly pine.
 ORGANISM Pinus taeda

REFERENCE 1 (bases 1 to 293)
 AUTHORS Sederoff, R.
 TITLE Molecular Basis of Wood Formation in the Pine Megagenome
 JOURNAL Unpublished (2000)
 COMMENT Contact: Johnson, Arthur
 North Carolina State University
 Tel: 919 515 7800
 Fax: 919 515 7801
 Email: ajohnson@unity.ncsu.edu
 Seq primer: T3.

FEATURES

source

Location/Qualifiers
 1..293
 /organism="Pinus taeda"
 /db_xref="taxon:3352"
 /clone="NXNV_073_C07"
 /clone_1lb="Nsf Xylem Normal wood Vertical"

/note="Vector: Bluescript SK; Site 1: Eco RI; The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGACAGC'."

BASE COUNT 72 a 67 c 60 g 83 t 11 others

ORIGIN

Query Match 100.0%; Score 13; DB 80; Length 293;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccatggaagcc 13
|||||

DB 93 CCCATGGAAGCC 105

RESULT 9
N28122 294 bp mRNA EST 25-JUL-1996
LOCUS M81422R Mouse brain, Stratagene Mus musculus cDNA 5' end, mRNA
DEFINITION
ACCESSION N28122
VERSION N28122.1 GI:1145978
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 294)
AUTHORS Beier, D. and Brady, K.
TITLE Mouse brain cDNAs
JOURNAL Unpublished (1995)
COMMENT On May 14, 1999 this sequence version replaced gi.4827793.
CONTACT: Sikeia JM
DEPARTMENT OF Pharmacology
UNIVERSITY OF Colorado Health Sciences Center
Box C236, 4200 E. 9th Ave, Denver CO 80262-0236
Tel: 3032708637
Fax: 3032707097
Email: nikkietally.uchsc.edu
Insert Length: 403 Std Error: 0.00
Seq primer: M13 Reverse.

FEATURES
source
1..294
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="Mouse brain, Stratagene"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP; Site 1: EcoR I; Site 2: Xho I; The mouse brain library (Stratagene) was constructed by oligo-(dT) priming and directional cloning in Uni-ZAP XR phage using whole brain mRNA from a Balb C post natal 20 day."

BASE COUNT 68 a 90 c 59 g 75 t 2 others

ORIGIN

Query Match 100.0%; Score 13; DB 86; Length 294;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccatggaagcc 13
|||||

DB 34 CCCATGGAAGCC 46

RESULT 10
A1352128 326 bp mRNA EST 13-FEB-1999
LOCUS q10f07.x1 Soares_total_fetus_NB2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:1940485 3', mRNA sequence.
ACCESSION A1352128

VERSION A1352128.1 GI:4089334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 326)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 888 Std Error: 0.00
Seq primer: -400P from Gldco
High quality sequence stop: 317.

FEATURES
source
1..326
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1940485"
/clone_id="Soares_total_fetus_NB2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGCGCGCTTAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 94 a 81 c 75 g 76 t

ORIGIN

Query Match 100.0%; Score 13; DB 38; Length 326;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccatggaagcc 13
|||||

DB 240 CCCATGGAAGCC 228

RESULT 11
A0411422 332 bp mRNA EST 16-FEB-2000
LOCUS fh12c03.y1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:296485 3', mRNA sequence.
DEFINITION
ACCESSION A0411422
VERSION A0411422.1 GI:6936963
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 332)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
JOURNAL On Apr 30, 1999 this sequence version replaced gi.4727490.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at:
www.bio.llnl.gov/dbfp/image/image.html
 Plate: LLAM64 row: F column: 6
 Seq primer: M13R1 reverse primer (AB1).

FEATURES

SOURCE

1. .332

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2964485"

/clone_lib="NIH-MGC.17"

/tissue_type="rhabdomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(C). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

75 a

70 c

46 g

141 t

ORIGIN

Query Match

100.0%; Score 13; DB 72; Length 332;

Best Local Similarity 100.0%; Pred. No. 8.8e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagcc 13

Db 171 CCCAATGGAAGCC 159

RESULT 12

LOCUS

DEFINITION

AA786244 338 bp mRNA EST 31-JUL-1998
 1711a1.f1 Aspergillus nidulans 24hr asexual developmental and vegetative CDNA lambda zap library Aspergillus nidulans CDNA clone 1711a1 3', mRNA sequence.

ACCESSION AA786244.1 GI:2846412

VERSION AA786244

KEYWORDS EST.

SOURCE Aspergillus nidulans.

ORGANISM Eukaryota; Fungi; Ascomycota; Eurotiiales; Trichocomaceae; anamorphic Trichocomaceae; Aspergillus.

REFERENCE 1 (bases 1 to 338)

AUTHORS Kupfer, D., Gray, J., Hauser, J., Lal, H., Martin, W., Aramayo, R., Prade, R. and Roe, B.

TITLE An Aspergillus nidulans EST Database

JOURNAL Unpublished (1998)

COMMENT Other ESTs: 1711a1.f1

Contact: Bruce A. Roe, University of Oklahoma, broegeu@ou.edu

Department of Chemistry and Biochemistry

Advanced Center for Genome Technology, University of Oklahoma

620 Parlington Oval, Norman, OK 73019, USA

Tel: 405 325 4912

Fax: 405 325 7762

Email: broegeu@ou.edu

We anticipate the future release of the CDNA clones to the Fungal Genetics Stock Center

Seq primer: M13-20

High quality sequence stop: 307.

Location/Qualifiers

1. .338

/organism="Aspergillus nidulans"

/strain="FGSC A26"

/db_xref="taxon:5072"

/clone="1711a1"

/clone_lib="Aspergillus nidulans 24hr asexual

developmental and vegetative CDNA lambda zap library"
 /tissue_type="vegetative mycelia, asexual structures"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; 5' end of CDNA cloned into EcoRI site of pBluescript 3' end of CDNA cloned into XhoI site of pBluescript"

BASE COUNT

92 a

86 c

68 g

91 t

1 others

ORIGIN

Query Match

100.0%; Score 13; DB 31; Length 338;

Best Local Similarity 100.0%; Pred. No. 8.8e+02;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cccaatggaagcc 13

Db 30 CCCAATGGAAGCC 42

RESULT 13

LOCUS

A1845473 338 bp mRNA EST 15-JUL-1999
 UI-M-AOI-aej-h-05-0-UI.a1 NIH-BMAP_MPG_N Mus musculus CDNA clone UI-M-AOI-aej-h-05-0-UI 3', mRNA sequence.

ACCESSION A1845473

VERSION A1845473.1 GI:5489379

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 338)

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mestrail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized pineal glands library cDNA library preparation: M.B.

Soares lab Clone distribution: NIH BMAP CDNA clones will be made available by the means that is soon to be determined. When NIH

determines the means for distribution of the BMAP CDNA clones, this record will be updated accordingly when that means is determined.

The following repetitive elements were found in this CDNA sequence:

1-30. >AT-richlow complexity

Seq primer: M13 forward

POLYA=yes.

Location/Qualifiers

1. .338

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-AOI-aej-h-05-0-UI"

/clone_lib="NIH-BMAP_MPG_N"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="Vector: p773D-pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The NIH-BMAP_MPG_N library is a normalized library constructed from mouse pineal gland. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Bonaldo, Lennon and Soares, Genome Research 6:

791-806, 1996. Tissue provided by Ms. Annie Novakovich,
Zivic-Miller Laboratories; TAG LIB=NIH_BMAP_MPG_N;
TAG_TISSUE=pineal-glands; TAG_SEQ=CAGAC"

BASE COUNT 86 a 78 c 84 g 90 t
ORIGIN
Query Match 100.0%; Score 13; DB 45; Length 338;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
|||||
Db 31 CCCAATGGAAGCC 43

RESULT 14
A1043778/c 348 bp mRNA EST 05-JUL-1999
LOCUS A1043778
DEFINITION UI-R-C0-jm-f-10-0-UI.s1 UI-R-C0 Rattus norvegicus cDNA clone
UI-R-C0-jm-f-10-0-UI 3', mRNA sequence.
A1043778
VERSION A1043778.1 GI:3290513
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 348)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
On Jun 5, 1998 this sequence version replaced gi:3189294.
CONTACT: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-dT track served to identify it as a clone from the normalized
adult 8-day-Embryo library. cDNA library Preparation: M. Fatima
Bonaldo, Ph.D. Clone distribution: clones will be available through
Research Genetics This clone is also available through the
I.M.A.G.E. Consortium at LNL (info@image.lnl.gov). IMAGE
ID-1783157
Seq primer: M13 Forward
POLYA-No.

FEATURES
source Location/Qualifiers
1. 348

/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C0-jm-f-10-0-UI"
/clone_1lb="UI-R-C0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-C0
library is a subtracted library derived from the UI-R-A1
and UI-R-E1 libraries. The UI-R-A1 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, and muscle. The UI-R-E1
library consisted of a mixture of individually tagged
normalized libraries constructed from 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which
allows identification of the library of origin of a clone

within the mixture. The subtracted library (UI-R-C0) was
constructed as follows: PCR amplified cDNA inserts from a
pool of UI-R-A1 and UI-R-E1 clones from which 3' ESTs had
been derived was used as a driver in a hybridization with
the pooled UI-R-A1 and UI-R-E1 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-C0
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)"

BASE COUNT 74 a 96 c 96 g 81 t 1 others
ORIGIN

Query Match 100.0%; Score 13; DB 34; Length 348;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
|||||
Db 248 CCCAATGGAAGCC 236

RESULT 15
T00893 350 bp mRNA EST 10-NOV-1992
LOCUS T00893
DEFINITION WEST01614 Early embryo, Strata gene (cat. #937007) Caenorhabditis
elegans cDNA clone CESTK50, mRNA sequence.
T00893
VERSION T00893.1 GI:277374
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Pelodermidae; Caenorhabditis.
1 (bases 1 to 350)
McCombie,W.R., Kelley,J.M., Aubin,L., Goscoechea,M.,
Fitzgerald,M.G., Wu,A., Adams,M.D., Dubnick,M., Kerlavage,A.R.,
Venter,J.C. and Fields,C.A.
Caenorhabditis elegans cDNAs
Unpublished (1993)
On May 5, 1995 this sequence version replaced gi:798322.
Other ESTs: WEST01615
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
Insert length: 1663 Std Error: 0.00
Seq primer: M13 Reverse
High quality sequence stop: 160.

TITLE Location/Qualifiers
1. 350
JOURNML
COMMENT /organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone="CESTK50"
/clone_1lb="Early embryo, Strata gene (cat. #937007)"

FEATURES
source Location/Qualifiers
1. 350

/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone="CESTK50"
/clone_1lb="Early embryo, Strata gene (cat. #937007)"

BASE COUNT 89 a 103 c 72 g 84 t 2 others
ORIGIN

Query Match 100.0%; Score 13; DB 89; Length 350;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cccaatggaagcc 13
|||||
Db 178 CCCAATGGAAGCC 190

Wed Sep 13 11:14:13 2000

us-09-542-718-5.rst

Page 9

Search completed: September 12, 2000, 22:47:00
Job time: 3802 sec

THIS PAGE BLANK (USPTO)